

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 06:49:22 ; Search time 310 Seconds
(without alignments)
467.855 Million cell updates/sec

Title: US-10-092-934-10
Perfect score: 2034
Sequence: 1 MEPSLLPLRLCNGAISAHK.....FIRGGVSPYLSGWSQTPDLR 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	3 AAB01399	Aab01399 Neuron-as
2	2034	100.0	375	5 ABB81538	Abb81538 Neural th
3	2034	100.0	375	5 AAE29142	Aae29142 AD7c-NTP
4	2034	100.0	375	6 ABR63268	Abr63268 AD7c-NTP
5	2034	100.0	375	6 ADA84017	Ada84017 Human POM
6	2034	100.0	375	6 ABO03024	Abu03024 Human neu
7	2034	100.0	375	6 ABB99774	Abb99774 Amino aci
8	2034	100.0	375	6 ABJ19445	Abj19445 AD7C-neur
9	2034	100.0	375	7 ADB37642	Adb37642 Human neu
10	2034	100.0	375	8 ADR14409	Adr14409 Human NF-
11	1415.5	69.6	397	2 AAR95913	Aar95913 Neural th
12	928	45.6	381	4 AAU30235	Aau30235 Novel hum
13	927.5	45.6	382	4 AAU32610	Aau32610 Novel hum
14	927.5	45.6	382	4 AAU31818	Aau31818 Novel hum
15	927.5	45.6	382	4 AAU32707	Aau32707 Novel hum
16	849	41.7	324	4 AAU29573	Aau29573 Novel hum
17	849	41.7	324	4 AAU29579	Aau29579 Novel hum
18	822.5	40.4	411	4 ABO08428	Abg08428 Novel hum
19	739	36.3	317	4 AAU33200	Aau33200 Novel hum
20	735	36.1	290	4 ABB21913	Abb21913 Novel hum
21	734.5	36.1	361	5 ABB68738	Abb68738 Human pro
22	729	35.8	449	5 ABB01057	Abb01057 Ovary cel
23	701.5	34.5	332	5 ADK34410	Adk34410 Novel hum
24	609	29.9	257	4 AAO10294	Aao10294 Human pol
25	590	29.0	276	4 ABB07919	Abb07919 Novel hum

RESULT 1
AAB01399
ID AAB01399 standard; protein; 375 AA.

AC AAB01399;
XX
DT 20-OCT-2000 (first entry)
DE Neuron-associated protein.

ALIGNMENTS

26	575	28.3	302	5 ADK36936	Adk36936 Novel hum
27	574	28.2	185	4 AAO06014	Aao06014 Human pol
28	569	28.0	241	4 ABB11464	Abb11464 Human neu
29	567.5	27.9	196	7 ADE09985	Ade09985 Novel pro
30	563	27.7	202	4 AAU30686	Aau30686 Novel hum
31	561	27.6	250	4 AAU31823	Aau31823 Novel hum
32	561	27.6	250	7 ADE09441	Ade09441 Novel pro
33	560	27.5	296	4 AAU33304	Aau33304 Novel hum
34	554.5	27.3	232	5 ADK36978	Adk36978 Novel hum
35	553	27.2	175	4 AAU31782	Aau31782 Novel hum
36	550.5	27.1	215	4 AAU31513	Aau31513 Novel hum
37	546	26.8	215	4 ABB07921	Abb07921 Novel hum
38	543.5	26.7	216	4 AAU32615	Aau32615 Novel hum
39	524	25.8	304	4 AAU30391	Aau30391 Novel hum
40	522.5	25.7	175	4 AAU31857	Aau31857 Novel hum
41	521.5	25.6	418	4 AAU31980	Aau31980 Novel hum
42	520	25.6	396	4 AAU30455	Aau30455 Novel hum
43	517	25.4	213	4 AAM25396	Aam25396 Human pro
44	504.5	24.8	196	4 AAU31786	Aau31786 Novel hum
45	503	24.7	189	4 AAU31834	Aau31834 Novel hum

Neuron associated protein; NEUAP; neurological disorder; epilepsy;
ischemic cerebrovascular disease; stroke; cerebral neoplasm;
Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
Parkinson's disease; demyelinating disease; meningitis; prion disease;
kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy;
muscular dystrophy; central nervous system; CNS;
peripheral nervous system; PNS; myopathy; schizophrenia;
actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;
cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
autoimmune disease; inflammation; acquired immunodeficiency syndrome;
AIDS; Addison's disease; adult respiratory distress syndrome; allergy;
ankylosing spondylitis; amyloidosis; anaemia; asthma;
Werner syndrome, trauma; human.

Homo sapiens.

WO200034477-A2.

15-JUN-2000.

10-DEC-1999; 99WO-US030408.

11-DEC-1998; 98US-00210083.

09-FEB-1999; 99US-0119365P.

16-MAR-1999; 99US-0124687P.

(INCY-) INCYTE PHARM INC.

Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-Young J;

Yang J, Lu DAM, Azinzai Y;

WPI; 2000-423423/36.

New human neuron-associated proteins and polynucleotides encoding them, useful for diagnosis, treatment and prevention of cell proliferative disorders including cancer, neuronal and neurological disorders.

PS Disclosure; Page 143-144; 145pp; English.

XX Human neuron-associated proteins (NEUPAP) can be used for treating or preventing a disorder associated with decreased expression or activity of NEUPAP. Antagonists of NEUPAP are useful for treating or preventing a disorder associated with increased expression or activity of NEUPAP. NEUPAP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUPAPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited, metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal parasitic, protozoal, and helminthic infections, and trauma. This protein was designated g3002527

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 3; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
 DB 1 MEFSLLLPRLECNCAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELTSDPSVSASQASARYTGHARLCLANFCGRNRVSLMCPSPWSPPELKQST 120
 DB 61 LHVGQAGLELTSDPSVSASQASARYTGHARLCLANFCGRNRVSLMCPSPWSPPELKQST 120

QY 121 CLSLPKCWYDRAAIVPGLFILFLHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 DB 121 CLSLPKCWYDRAAIVPGLFILFLHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180

QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240

QY 241 WDYRRPRLANFFVFLVEMGTMFARLILISGPCDLPASASQASAGITGVSHARLILNFC 300
 DB 241 WDYRRPRLANFFVFLVEMGTMFARLILISGPCDLPASASQASAGITGVSHARLILNFC 300

QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 360
 DB 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 360

QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 2
 ABB81538 standard; protein; 375 AA.
 XX ABB81538;
 AC ABB81538;
 XX 02-SEP-2002 (first entry)
 DT
 XX

DE Neural thread protein (NTP).

XX Neural thread protein; NTP; Harlil peptide; Alzheimer's disease;
 KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
 KW hypoxia; ischaemia; cerebral infarction.
 XX Homo sapiens.
 OS W0200234915-A2.
 PN 02-MAY-2002.
 PD 25-OCT-2001; 2001WO-US042813.
 PF 27-OCT-2000; 2000US-00697590.
 PR (NYMO-) NYMOX PHARM CORP.
 FA Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
 XX WPI: 2002-507998/54.
 XX N-PSDB; ABN89470.
 DR New Harlil peptide sequences of the Neural Thread Protein, useful in therapeutic assays, e.g. as targets for developing drugs for treating Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing these diseases.

XX Example 1; Fig 1; 53pp; English.

XX The present invention describes a neural thread protein (NTP) peptide having an amino acid sequence selected from ABB81511 to ABB81529 and their homologues, which are referred collectively as Harlil peptides (I). (I) have neuroprotective, nootropic, vasotropic and cerebroprotective activities, and can be used in peptide therapy. The Harlil peptide sequences can be used as analogues for NTP in therapeutic or diagnostic assays by replacing NTP with the peptide in such an assay. The Harlil peptides are also useful as a trap material in a diagnostic or therapeutic assay. Therefore, the Harlil peptides are useful in binding assays, protein and antibody purification, therapeutics or diagnostics. In particular, the peptides are also useful for diagnosing Alzheimer's disease, Down's syndrome, neuroectodermal tumours, astrocytomas, glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides are also useful as targets for drug development for the treatment of these diseases. The present sequence represents a neural thread protein given in the exemplification of the present invention

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 5; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
 DB 1 MEFSLLLPRLECNCAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELTSDPSVSASQASARYTGHARLCLANFCGRNRVSLMCPSPWSPPELKQST 120
 DB 61 LHVGQAGLELTSDPSVSASQASARYTGHARLCLANFCGRNRVSLMCPSPWSPPELKQST 120

QY 121 CLSLPKCWYDRAAIVPGLFILFLHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 DB 121 CLSLPKCWYDRAAIVPGLFILFLHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180

QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240

QY 241 WDYRRPRLANFFVFLVEMGTMFARLILISGPCDLPASASQASAGITGVSHARLILNFC 300
 DB 241 WDYRRPRLANFFVFLVEMGTMFARLILISGPCDLPASASQASAGITGVSHARLILNFC 300

QY 301 LFEMESHVSVTQAGVQVWPNLGLSLQPLPGLKRFKFSCLSLPSSWDYGHLLPPHPANFCIFIRGG 360
Db 301 LFEMESHVSVTQAGVQVWPNLGLSLQPLPGLKRFKFSCLSLPSSWDYGHLLPPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 3
AAE29142
ID AAE29142 standard; protein; 375 AA.
XX
AC AAE29142;
XX
DT 27-JAN-2003 (first entry)
XX
DE AD7c-NTP protein.
XX

KW Neural thread protein; NTP; hyperplasia; hypertrophy; arteriosclerosis;
KW haemorrhoid; gene therapy; tumour; vascular disease; atherosclerosis;
KW inflammatory disease; nutritional deficiency disease; genetic disease;
KW autoimmune disease; metabolic disease; traumatic disease; intoxication;
KW infectious disease; congenital malformation; enzyme deficiency disease;
KW amyloid disease; fibrosis disease; storage disease; radiation disease;
KW poisoning; environmental disease; endocrine disease; protein therapy;
KW degenerative disease; mechanical disease; AD7c-NTP protein.
XX
OS Unidentified.
XX
PN W0200274323-A2.
XX
PD 26-SEP-2002.
XX
XX 08-MAR-2002; 2002WO-IB001959.
XX
PR 08-MAR-2001; 2001US-0273957P.
XX
XX (AVER/) AVERBACK P.
XX
XX Averbach P;
XX
XX WPI; 2002-759864/82.
DR N-PSDS; AAD46671.
XX
XX

Treating a condition in a patient requiring removal or destruction of
cells, such as a benign or malignant tumor of a tissue or an inflammatory
disease, comprises administering a neural thread protein (NTP) or a NTP
gene to a mammal.

Example 2; Fig 1; 70pp; English.

The invention relates to a method for treating a condition in a patient
requiring removal or destruction of cells. The method involves
administering to a mammal a neural thread protein (NTP), or administering
to a tumour or other target cell a NTP gene, where the expression of the
NTP gene is induced resulting in expression of the NTP protein. The
method and NTP are useful for treating a condition in a patient requiring
removal or destruction of cells, such as a benign or malignant tumour of
a tissue, a hyperplasia, hypertrophy, or overgrowth of a tissue,
preferably tonsillar hypertrophy or prostatic hyperplasia, a virally,
bacterially, or parasitically altered tissue, or a malformation of a
tissue. Other conditions include a cosmetic modification to a tissue,
such as removal of unwanted facial hair, warts or unwanted fatty tissue,
a vascular disease, particularly atherosclerosis or arteriosclerosis,
haemorrhoids, or varicose veins, an inflammatory disease, autoimmune
disease, metabolic disease, hereditary/genetic disease, traumatic disease,
or physical injury, nutritional deficiency disease, infectious disease,
congenital malformation, amyloid disease, fibrosis disease, storage
disease, enzyme deficiency disease, poisoning, intoxication, degenerative
disease, radiation disease, environmental disease, endocrine disease or
mechanical disease. The invention is useful in protein therapy and gene
therapy. The present sequence is AD7c-NTP protein

XX SQ Sequence 375 AA;
Query Match 100.0%; Score 2034; DB 5; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-198;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEPSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MEPSLLLPRLCNGAISAHRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
QY 61 LHVGQAGLELPTSDDFSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDDFSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
QY 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLQSTPEIKHPPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLQSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQVWPNLGLSLQPLPGLKRFKFSCLSLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQVWPNLGLSLQPLPGLKRFKFSCLSLSS 240
QY 241 WDYRRPRLANPFFVLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANPFFVLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSVTQAGVQVWPNLGLSLQPLPGLKRFKFSCLSLPSSWDYGHLLPPHPANFCIFIRGG 360
Db 301 LFEMESHVSVTQAGVQVWPNLGLSLQPLPGLKRFKFSCLSLPSSWDYGHLLPPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 4
ABR63268
ID ABR63268 standard; protein; 375 AA.
XX
AC ABR63268;
XX
DT 28-AUG-2003 (first entry)
XX
DE AD7c-NTP protein.
XX
KW Cytostatic; Antibacterial; Immunosuppressive; Antiinflammatory;
KW neural thread protein; NTP; tumour.
XX
OS Unidentified.
XX
PN W02003008443-A2.
XX
PD 30-JAN-2003.
XX
XX 19-JUL-2002; 2002WO-CA001105.
XX
XX 19-JUL-2001; 2001US-0306150P.
PR 19-JUL-2001; 2001US-0306161P.
PR 16-NOV-2001; 2001US-0331477P.
XX
XX (NYMO-) NYMOX CORP.
XX
XX Averbach PA;
XX
XX WPI; 2003-247999/24.
XX
XX Novel neural thread protein peptide, referred as cell death peptide,
XX useful for treating prostatic hyperplasia, psoriasis, eczema, dermatosis,
XX atherosclerosis, cosmetic modification to skin, throat, mouth, muscle.
XX
XX Disclosure; Fig 1; 77pp; English.
XX

CC The present invention relates to a neural thread protein (NTP) peptide
 CC referred to as cell death peptide. Thought to be cytostatic,
 CC antibacterial, immunosuppressive and antiinflammatory. It is useful for
 CC treating a condition in a patient requiring removal or destruction of
 CC cells, for treating a condition such as benign or malignant tumor,
 CC inflammatory disease, autoimmune disease and infectious disease. The
 CC peptide useful for treatment is derived from the amino acid sequence for
 CC a pancreatic thread protein. The peptide is conjugated, linked or bound
 CC to a molecule chosen from antibody or its fragment, antibody-like binding
 CC molecule, where the molecule has a higher affinity for binding to a tumor
 CC or other target than binding to other cells. Treatment using NTP peptides
 CC can remove benign tumors with less risk and fewer of the undesirable side
 CC effects of surgery. The present sequence is an NTP amino acid sequence
 XX
 SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEFSLLPRLECNCAISAHNRLRLPGSSDSPASASPVAGITGMCTHARLIILYFFLVEMEF 60
 DB 1 MEFSLLPRLECNCAISAHNRLRLPGSSDSPASASPVAGITGMCTHARLIILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPPSVASQSARVYRTGHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPPSVASQSARVYRTGHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKCWYDRAAIVPGLFILFLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 DB 121 CLSLPKCWYDRAAIVPGLFILFLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFNLFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFNLFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 QY 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHARLIENFC 300
 DB 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHARLIENFC 300
 QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 DB 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 5
 ADA84017
 ID ADA84017 standard; protein; 375 AA.
 XX
 AC ADA84017;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human POM80 protein.
 XX
 KW human; marker; expressed sequence tag; EST; arabidopsis; tumour;
 KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2002103028-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 30-MAY-2002; 2002WO-IB004189.
 XX
 PR 30-MAY-2001; 2001US-0293999P.
 PR 22-OCT-2001; 2001US-0330457P.

PR 19-FEB-2002; 2002US-0357144P.
 XX (BIOM-) BIOMEDICAL CENT.
 XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
 XX WPI; 2003-175241/17.
 DR N-PSDB; ADA84016.
 XX
 PT Determining if a nucleic acid is a marker for a phenotype/cell type of
 PT interest, by global comparison of expressed sequence tags known to be
 PT expressed in the phenotype/cell type with all ESTs expressed in normal
 PT tissue.
 XX
 PS Claim 29; Page 398-400; 516pp; English.
 XX
 CC The invention relates to a novel method for determining if a nucleic acid
 CC is a marker for a predetermined phenotype/cell type of interest from a
 CC biological species. The method comprises performing a global comparison
 CC of a group of expressed sequence tags (ESTs) known to be expressed in the
 CC phenotype/cell type of interest with all ESTs expressed in normal tissue
 CC in order to identify ESTs that are preferentially expressed in the
 CC phenotype/cell of interest. A method of the invention is useful for
 CC determining whether a nucleic acid is a marker for a predetermined
 CC phenotype or cell type of interest from a biological species, preferably
 CC Arabidopsis or human. The cell type of interest is an abnormal cell such
 CC as a tumour cell, and the predetermined phenotype is a stress-induced
 CC phenotype such as hyperosmotic stress or high salt conditions. A method
 CC of the invention is also useful for determining the progression of colon
 CC cancer in a human, for detecting a tumour cell, and for regulating or
 CC preventing the growth of a tumour cell. An antibody of the invention is
 CC useful for detecting the absence or presence of peptides encoded by
 CC tumour-associated markers. A polypeptide of the invention is useful as an
 CC immunogen for vaccinating an animal. The present sequence represents a
 CC tumour-associated antigen of the invention.
 XX
 SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEFSLLPRLECNCAISAHNRLRLPGSSDSPASASPVAGITGMCTHARLIILYFFLVEMEF 60
 DB 1 MEFSLLPRLECNCAISAHNRLRLPGSSDSPASASPVAGITGMCTHARLIILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPPSVASQSARVYRTGHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPPSVASQSARVYRTGHARLCLANFCGRNVRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKCWYDRAAIVPGLFILFLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 DB 121 CLSLPKCWYDRAAIVPGLFILFLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFNLFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 DB 181 VAGTKDMHYTWLIFIFNLFRLQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
 QY 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHARLIENFC 300
 DB 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAGITGVSHARLIENFC 300
 QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 DB 301 LFEMESHVSTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 6
 ABU03024

AD	ABU03024 standard; protein; 375 AA.	SQ	Sequence 375 AA;
AC	ABU03024;		Query Match 100.0%; Score 2034; DB 6; Length 375;
XX			Best Local Similarity 100.0%; Pred. No. 1e-198;
XX			Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DT	20-JAN-2003 (first entry)		
DE	Human neural thread protein AD7C-NTP.		
XX		QY	1 MEFSLLPLRLCNGAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFLVEMEF 60
KW	Neural thread protein; NTP-peptide; AD7C-NTP; surgical excision;	DB	1 MEFSLLPLRLCNGAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFLVEMEF 60
KW	transplantation; grafting; chemotherapy; immunotherapy; vaccination;		
KW	ablation; cryotherapy; laser therapy; phototherapy; gene therapy;	QY	61 LHVGQAGLELPTSDSDSVASQASRYRTGHARLCLANFCGRNVRVSLMCPSPWPELKQST 120
KW	radiation; tumor; hyperplasia; hypertrophy; overgrowth of tissue;	DB	61 LHVGQAGLELPTSDSDSVASQASRYRTGHARLCLANFCGRNVRVSLMCPSPWPELKQST 120
KW	malformation of tissue; tonsillary hypertrophy; prostatic hyperplasia;		
KW	cosmetic modification; vascular disease; atherosclerosis;	QY	121 CLSLPKCWDYRRRAVFGFLFELFRLHRCPTLTQDEVQWCDHSSLQSPSTPIKHPPASASQ 180
KW	arteriosclerosis; haemorrhoid; varicose vein; inflammatory disease;	DB	121 CLSLPKCWDYRRRAVFGFLFELFRLHRCPTLTQDEVQWCDHSSLQSPSTPIKHPPASASQ 180
KW	autoimmune disease; metabolic disease; traumatic disease;		
KW	physical injury; nutritional deficiency disease; infectious disease;	QY	181 VAGTKDHHYTWLIFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
KW	amyloid disease; fibrosis disease; storage disease; poisoning;	DB	181 VAGTKDHHYTWLIFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
KW	congenital malformation; enzyme deficiency disease; radiation disease;		
KW	intoxication; environmental disease; endocrine disease; degenerative disease; mechanical disease.	QY	241 WDYRRPRLANFFVFLVEMGFTWFARLILISGCDLPASASQAGITGVSHHARLIENFC 300
OS	Homo sapiens.	DB	241 WDYRRPRLANFFVFLVEMGFTWFARLILISGCDLPASASQAGITGVSHHARLIENFC 300
XX			
XX	WO200297030-A2.		
XX			
XX	05-DEC-2002.		
XX			
XX	24-MAY-2002; 2002WO-CA000759.		
XX			
XX	25-MAY-2001; 2001US-0293156P.		
PR	(NYMO-) NYMOX CORP.		
PA	Averback PA;		
XX			
XX	WPI; 2003-041406/03.		
DR			
XX			
XX	Novel peptides similar in amino acid sequence to neural thread proteins		
PT	(NTP), useful for treating unwanted cellular proliferations such as		
PT	malignant tumors and prostatic hyperplasia.		
XX			
XX	Disclosure; Fig 1; 78pp; English.		
PS			
XX			
CC	The invention describes an NTP-peptide (I) comprising at least one amino		
CC	acid sequence corresponding to part of the amino acid sequence of a		
CC	neural thread protein, AD7C-NTP. The invention provides a method of		
CC	treating a condition requiring removal or destruction of cells of a		
CC	mammal comprising administering to a mammal, a therapeutic amount of (I).		
CC	The treatment is administered to the mammal before, during or after		
CC	surgical excision, transplantation, grafting, chemotherapy,		
CC	immunotherapy, vaccination, thermal or electrical ablation, cryotherapy,		
CC	laser therapy, phototherapy, gene therapy and/or radiation. The method is		
CC	useful for treatment of benign or malignant tumour; hyperplasia,		
CC	hyper trophy or overgrowth of tissue; virally, bacterially or		
CC	parasitically altered tissue; malformation of tissue selected from lung,		
CC	breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney,		
CC	sinus, colon, intestine, rectum, esophagus, heart, spleen, salivary		
CC	gland, blood, brain and its coverings, spinal cord, muscle, connective		
CC	tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary,		
CC	reproductive organs, liver, hair, gall bladder, eye, ear, nose, throat,		
CC	tonsils, mouth and lymph nodes and lymphoid system; tonsillary		
CC	hyper trophy; prostatic hyperplasia; cosmetic modification to a tissue;		
CC	vascular disease (atherosclerosis or arteriosclerosis); haemorrhoids;		
CC	varicose veins; inflammatory disease; autoimmune disease; metabolic		
CC	disease; hereditary/genetic disease; traumatic disease; physical injury;		
CC	nutritional deficiency disease; infectious disease; amyloid disease;		
CC	fibrosis disease; storage disease; congenital malformation; enzyme		
CC	deficiency disease; poisoning; intoxication; environmental disease;		
CC	radiation disease; endocrine disease; degenerative disease and mechanical		
CC	disease. This is the amino acid sequence of the human neural thread		
XX	protein AD7C-NTP		
XX			

RESULT 7	
ABB99774	
ID	ABB99774 standard; protein; 375 AA.
XX	
AC	ABB99774;
XX	
DT	24-MAR-2003 (first entry)
XX	
DE	Amino acid sequence of human neuronal thread protein AD7c-NTP.
XX	
KW	Human; neuronal thread protein; AD7c-NTP; Alzheimer's disease; histone;
KW	neurodegeneration; in vivo gene expression; amphipathic compound;
XX	gene therapy.
OS	Homo sapiens.
XX	
PN	WO200299036-A2.
XX	
PD	12-DEC-2002.
XX	
PF	28-MAY-2002; 2002WO-US016429.
XX	
PR	01-JUN-2001; 2001US-00872968.
XX	
PA	(RHOD-) RHODE ISLAND HOSPITAL.
XX	
PI	Wands JR, De La Monte SM;
XX	
DR	WPI; 2003-140605/13.
XX	
XX	N-PSDB; ABZ23236.
PT	Inducing prolonged in vivo gene expression in mammal by contacting
PT	neuronal tissue with composition comprising Alzheimer's disease-
PT	associated neural thread protein 7c antisense nucleic acid, histone,
PT	amphipathic compound.
XX	
PS	Disclosure; Page 35; 69pp; English.
XX	

CC The present sequence represents a human neuronal thread protein AD7c-NTP.
 CC AD7c-NTP is overexpressed in brains with Alzheimer's disease at early and
 CC intermediate stages of neurodegeneration. The expression of AD7c-NTP may
 CC be reduced using the method of the invention. The specification describes
 CC a method for inducing prolonged in vivo gene expression in a mammal. The
 CC method comprises contacting a non-muscular tissue with a composition
 CC comprising a nucleic acid, histone and an amphipathic compound. The
 CC method is useful for inducing prolonged in vivo gene expression in non-
 CC muscular tissue of a mammal, e.g. neuronal tissue, central nervous system
 CC (CNS) tissue, tissue comprising a post-mitotic neuronal cell, cortical
 CC neuronal cell or hippocampal neuronal cell, glial cell, or vascular
 CC endothelial cell. The method is useful in gene therapy applications to
 CC treat Alzheimer's disease, where the composition comprises antisense
 CC AD7c-NTP nucleic acid

XX Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
 DB 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKWDYRRAAVPGFLIFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 180
 DB 121 CLSLPKWDYRRAAVPGFLIFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLFIFFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 240
 DB 181 VAGTKDMHYTWLFIFFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 240
 QY 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 DB 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 QY 301 LFEMESHVTVQAGVQWPNLGSLLQPLPGLKRFSCLSLSPSSWDYGHLPHPANFCIFIRGG 360
 DB 301 LFEMESHVTVQAGVQWPNLGSLLQPLPGLKRFSCLSLSPSSWDYGHLPHPANFCIFIRGG 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 8
 ABJ19445
 ID ABJ19445 standard; protein; 375 AA.

XX AC ABJ19445;

XX DT 27-MAR-2003 (first entry)

XX DE AD7C-neural thread protein.

XX KW Nootropic; neuroprotective; cell death; tissue necrosis; NTP;
 KW neural thread protein; neurodegenerative disorder; Alzheimer's disease.

XX OS Unidentified.

XX FN WO200292115-A2.

XX PD 21-NOV-2002.

XX PF 16-MAY-2002; 2002WO-CA000712.

XX PR 16-MAY-2001; 2001US-0290971P.

XX

PA (NYMO-) NYMOX CORP.

XX PI Averbach PA;

XX DR WPI; 2003-129234/12.

XX DR N-PSDB; AAL54224.

XX PT Preventing and/or inhibiting cell death and/or tissue necrosis in a
 PT tissue for treating a neurodegenerative disorder, e.g. Alzheimer's
 PT disease, by contacting the live tissue with at least one segment of
 PT neural thread proteins (NTP).

XX PS Disclosure; Fig 1; 60pp; English.

XX CC The invention relates to a novel method for preventing and/or inhibiting
 CC cell death and/or tissue necrosis in a tissue comprising contacting the
 CC live tissue with at least one segment of neural thread proteins (NTP).
 CC The methods are composition are useful for treating a neurodegenerative
 CC disorder, such as Alzheimer's disease. This sequence represents the AD7C-
 CC NTP protein relating to the invention

XX SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
 DB 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
 QY 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 DB 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
 QY 121 CLSLPKWDYRRAAVPGFLIFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 180
 DB 121 CLSLPKWDYRRAAVPGFLIFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLFIFFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 240
 DB 181 VAGTKDMHYTWLFIFFLFLHRCPRTLQDEVQWCDHSSLPSTPEIKHPPASASQ 240
 QY 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 DB 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIFNFC 300
 QY 301 LFEMESHVTVQAGVQWPNLGSLLQPLPGLKRFSCLSLSPSSWDYGHLPHPANFCIFIRGG 360
 DB 301 LFEMESHVTVQAGVQWPNLGSLLQPLPGLKRFSCLSLSPSSWDYGHLPHPANFCIFIRGG 360
 QY 361 VSPYLSGWSQTPDLR 375
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 9
 ADB37642

ID ADB37642 standard; protein; 375 AA.

XX AC ADB37642;

XX DT 04-DEC-2003 (first entry)

XX DE Human neural thread protein AD7c-NTP.

XX KW Cytostatic; Antitumour; Antipsoriatic; Dermatological;
 KW Antiatherosclerotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory;
 KW Immunosuppressive; Tranquillizer; Antiemetic; Virucide; AD7c-NTP;
 KW neural thread protein; neuritic sprouting.

XX OS Homo sapiens.

XX

PN WO2003008444-A2.
 XX 30-JAN-2003.
 XX 19-JUL-2002; 2002WO-CA001106.
 XX 19-JUL-2001; 2001US-0306150P.
 PR 19-JUL-2001; 2001US-0306161P.
 PR 16-NOV-2001; 2001US-0331477P.
 XX (NYMO-) NYMOX CORP.
 XX
 XX Averbach PA, Gemmell J;
 XX WPI; 2003-248000/24.
 DR N-PSDB; ADB37519.
 XX
 PT Novel Related peptide or AD7c-neural thread peptide, useful for treating
 PT unwanted cellular proliferations, glandular hyperplasia, unwanted facial
 PT hair, warts and unwanted fatty tissue.
 XX
 PS Disclosure; Fig 1; 10pp; English.
 XX
 CC The present invention relates to AD7c-neural thread protein (NTP) and
 CC related proteins and peptides (I; ADB37528-ADB37641). The sequences are
 CC useful for treating a condition in a patient requiring removal or
 CC destruction of cells. The condition can be selected from benign or
 CC malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a
 CC tissue, virally, bacterially or parasitically altered tissue, or
 CC malformation of a tissue, where the tissue is selected from lung, breast,
 CC stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, salivary
 CC colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary
 CC gland, blood, brain and its coverings, spinal cord and its coverings.
 CC muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,
 CC pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,
 CC throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is
 CC preferably tonsillary hypertrophy, prostatic hyperplasia, psoriasis,
 CC eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,
 CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular
 CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose
 CC veins, inflammatory disease, autoimmune disease, metabolic disease,
 CC hereditary/genetic disease, traumatic disease or physical injury,
 CC nutritional deficiency disease, infectious disease, amyloid disease,
 CC fibrosis disease, storage disease, congenital malformation, enzyme
 CC deficiency disease, poisoning, intoxication, environmental disease,
 CC radiation disease, endocrine disease, degenerative disease and mechanical
 CC disease. The peptides are useful for treating unwanted cellular
 CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial
 CC hair, warts and unwanted fatty tissue, or for preparing antibodies that
 CC recognize and/or bind to Related proteins, Related peptides or NTP
 CC peptides. The present sequence is AD7c-NTP, used to illustrate the
 CC invention. AD7c-NTP is a -41kd membrane associated phosphoprotein with
 CC functions associated with neuritic sprouting.
 XX
 SQ Sequence 375 AA;
 Query Match 100.0%; Score 2034; DB 7; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1e-198;
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEFSLLLPRLCNCAISAHNRNLPGSSDSPASASPVAGITGMCTHARLLIYPLVMEF 60
 DB 1 MEFSLLLPRLCNCAISAHNRNLPGSSDSPASASPVAGITGMCTHARLLIYPLVMEF 60
 QY 61 LHVQAGLELPTSDPVSASQSAARYTGHARLCLANFCGRNVRSLMCPSPWSPQLQST 120
 DB 61 LHVQAGLELPTSDPVSASQSAARYTGHARLCLANFCGRNVRSLMCPSPWSPQLQST 120
 QY 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 DB 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
 QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240

Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
 QY 241 WDYRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
 Db 241 WDYRRPRLANFVFLVEMGFTMFARLLISGPCDLPASASQAGITGVSHHARLIFNFC 300
 QY 301 LFEMESHVTVQAGVQWRNLGSLQPLPPGKLFSCPSCLSLPSSWDYGHLPHPANFCIFIRGG 360
 Db 301 LFEMESHVTVQAGVQWRNLGSLQPLPPGKLFSCPSCLSLPSSWDYGHLPHPANFCIFIRGG 360
 QY 361 VSPYLSGWSQTPDLR 375
 Db 361 VSPYLSGWSQTPDLR 375

RESULT 10
 ADR14409
 ID ADR14409 standard; protein; 375 AA.
 XX
 AC ADR14409;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated protein SeqID410.
 XX
 KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
 KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnery; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IGM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004WO-US000798.
 XX
 PR 14-JAN-2003; 2003US-0440068P.
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Nadler SG, Neubauer MG, Feder JN, Carman J;
 PI WPI; 2004-562168/54.
 DR N-PSDB; ADR14408.
 DR
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 PS Claim 6; SEQ ID NO 410; 237pp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
 CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,

CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnery activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, an inflammatory disorder
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked adhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human protein which
 CC is subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

XX Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 8; Length 375;

Best Local Similarity 100.0%; Pred. No. 1e-198;

Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60

DB 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARCLANFCGRNVRSLMCPSPKQST 120

DB 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARCLANFCGRNVRSLMCPSPKQST 120

QY 121 CLSLPKCDYRRRAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180

DB 121 CLSLPKCDYRRRAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180

QY 181 VAGTKDMHHYTWLIFIFTNFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

DB 181 VAGTKDMHHYTWLIFIFTNFLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

QY 241 WDYRPPRLANFFVFLVEMGFTMFARLLISGPCDLPASASQSAAGITGVSHHARLIENFC 300

DB 241 WDYRPPRLANFFVFLVEMGFTMFARLLISGPCDLPASASQSAAGITGVSHHARLIENFC 300

QY 301 LFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

DB 301 LFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPILSGWSQTPDLR 375

DB 361 VSPILSGWSQTPDLR 375

RESULT 11

AAU30235

ID AAU30235 standard; protein; 397 AA.

XX AAU30235;

AC AAU30235;

XX 13-NOV-1996 (first entry)

DT 13-NOV-1996 (first entry)

DE Neural thread protein.

XX Neural thread protein.

XX Neural thread protein; NTP; diagnosis; detection; Alzheimer's disease;

KW neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;

KW binding fragment.

XX Homo sapiens.

OS Homo sapiens.

XX WO9615272-A1.

PN 23-MAY-1996.

PD 14-NOV-1995; 95WO-US017111.

XX 14-NOV-1995; 94US-00340426.

XX 14-NOV-1994; 94US-00340426.

XX (GEO) GEN HOSPITAL CORP.

PA De La Monte S, Wands JR;

XX WPI; 1996-259865/26.

XX N-PSDB; AAT27738.

XX Detection of neural thread protein in diagnosis of Alzheimer's disease -

PT also NTP DNA and protein sequences used in gene and anti-sense therapy.

XX Claim 22; Page 171-172; 238pp; English.

XX A method for detecting the presence of neural thread protein (NTP) having

CC a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject

CC comprises (a) contacting a sample from a human subject that is suspected

CC of containing the NTP with at least one molecule capable of binding to

CC the protein; and (b) detecting any of the molecule bound to the protein.

CC The binding molecule is selected from an antibody free of natural

CC impurities, a monoclonal antibody or a binding fragment of either of

CC these. The method may be used for diagnosing the presence of Alzheimer's

CC disease, neuroectodermal tumours and a malignant astrocytoma in a human

XX Sequence 397 AA;

Query Match 69.6%; Score 1415.5; DB 2; Length 397;

Best Local Similarity 74.6%; Pred. No. 1.7e-135;

Matches 285; Conservative 10; Mismatches 44; Indels 43; Gaps 7;

QY 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60

DB 1 MEFSLLLPRLCNGAISAHRNLRPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLEL-----PTSDPSVSASQSAARYTGHARCLANFCGRNVRSLMCP 110

DB 61 LHVGQAGLELRLPQMLPSRPPKVLDTGLSTMPG-----LCLANFCGRNVRSLMCP 110

QY 111 SWSPKELKQSTCLSLPKCDYRRRAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPE 170

DB 111 SWSPKELKQSTCLSLPKCDYRRRAVPGLFILFFLRHRCPTLTQDEVQWCDHSSLPQSTPE 170

QY 171 ----IKHPPASASQVAGTKDMHHYTWLIFIFTNFLRQSLNSVTQAGVQWRNLGSLQPLP 226

DB 171 SSILLPQPP---KVAGTKDMHHYTWLIFIFTNFLRQSLNSVTQAGVQWRNLGSLQPLP 226

QY 227 PGFKLFSCPSLLSSWDYRRPPRLANFPVFLVEMGF-----TWPAR----LILISGCDLPA 278

DB 227 PGFKLFSCPSLLSSWDYRRPPRLANFPVFLVEMGF-----TWPAR----LILISGCDLPA 278

QY 279 SASQAGITGVSHHARLIENFCLFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLP 338

DB 282 --PKVGLQDVTPPTARPIFNFCFLFEMESHVTVQGVQWPNLGSLOPLPGPKLRFSCLSLP 339

QY 339 SSWDYGHLPPHPANFCIFIRGG 360

DB 340 SSWDYGHLPPHPANFCIFIRGG 361

RESULT 12

AAU30235

ID AAU30235 standard; protein; 381 AA.

XX AAU30235;

AC AAU30235;

XX 18-DEC-2001 (first entry)

DT 18-DEC-2001 (first entry)

XX Novel human secreted protein #726.
DE Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
XX WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 264-265; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX SQ Sequence 381 AA;
Query Match 45.6%; Score 928; DB 4; Length 381;
Best Local Similarity 55.9%; Pred. No. 1e-85;
Matches 223; Conservative 27; Mismatches 101; Indels 48; Gaps 14;
Qy 4 SLLPLRLECNAGISAHRLNLPDSSSPASAPVAGITGCTHARLILYFLVEMEFHV 63
Db 4 SLLPLRLECNAGILAHCLNCLSGSSDPASASQVTGKCHHTQLI-FVFLVEMGFPHI 62
Qy 64 QAQGLELPTSDPSVASQSRVYTG--HARLCLANFCGNRVLNCPGW-----112
Db 63 QAQGLELPTSDPTL-ASQSAGI-TGVNHHAWLFF--FCSRDTVSLCPGWSRVAXSRIT 118
Qy 113 --SPELKGSTCLSLPKCWDYERAAV-PGLTILFLRHCPTLTQDEVQWCDHSSLQP---166
Db 119 ATSPGLKXFACTSLPSSRDYRHPHPGNCFCIF-----GRDEVSPC-----WFGWF 164
Qy 167 STBIKHPPASASQVAGTKMHHYTW----LIFIFINFLRQSLNSVTQAGVQWRNLGSL 222
Db 165 XTPDLRPPASASQSAEIIIGVSHHTWQEVFLNLFILYRLWSLDSVAQARVQRDLGSL 224
Qy 223 QPLPPGFKLFCPSLLGSWYRR--PPRIANFVFLVEMGFMTFARLILISGPCDLPASAS 281
Db 225 QAPPFRKPKFCLSLPSSWDYRRPPHPANFVFLVETGFTVLARRVLISXPRDLPASAS 284

Qy 282 QSAGITGVSHHARLIFNFCLEFEMESHVSVTQAGVQWPNLQSLQPLPGLKRFCLSLPSSW 341
Db 285 QSAGITGVSHHARLIFNFCLEFEMESHVSVTQAGVQWPNLQSLQPLPGLKRFCLSLPSSW 342
Qy 342 DYGHLPPEHPAN-----FCIFIRGGVSPYLSGWSQTPDLR 375
Db 343 DYRCITPPCPNPLFIYLFYFHRDEGSLCCPCGWSXTPELK 381
RESULT 13
AAU32610
ID AAU32610 standard; protein; 382 AA.
XX AC AAU32610;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #3101.
XX KW Human; vaccination; gene therapy; nutritional supplement;
XX KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US008656.
XX PR 18-APR-2000; 2000US-00552929.
XX PR 26-JAN-2001; 2001US-00770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 637; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX SQ Sequence 382 AA;
Query Match 45.6%; Score 927.5; DB 4; Length 382;
Best Local Similarity 55.8%; Pred. No. 1.1e-85;
Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;
Qy 4 SLLPLRLECNAGISAHRLNLPDSSSPASAPVAGITGCTHARLILYFLVEMEFHV 63
Db 4 SLLPLRLECNAGILAHCLNCLSGSSDPASASQVTGKCHHTQLI-FVFLVEMGFPHI 62

```
QY 64 GQAGLELPTSDPSVSASQSYRTG--HHARLCLANFCGRNRVSLMCPWS----- 113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 AQAGLELLTSDSPTL-ASQSAGI-TGVNHAWLFF--FCSRDTVSLCPGWSRVAXSRIT 118
QY 114 ----PELKQSTCLSLPKCWDYRRAV-PGLFILFPLRHRCPTLTQDEVQVCHSSLP-- 166
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 ATAPGLKXKFCFLSPSSRDYRHVPHPGNCIF-----GRDEVSPC-----WPGW 164
QY 167 -STPEIKHPPASQVAGTKDMHYTW----LIFIFIFNFRQSLNSVTQAGVQWRNLGS 221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 FXTPLRLYPASASQSAEIIIGVSHHTWQEVFLFNLFYLRWLSDSVAQARVQRDLGS 224
QY 222 LQPLPPGKLFSCPSLLSSWDYRR--PPRLANFFVFLVEMGMTFARLILISGPCDL PASA 280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 LQAPPFRKPFCLSLPSSWDYRRPPPPHPPANFFVFLVETGTVLARRVLISXPRDL PASA 284
QY 281 SOSAGITGVSHHARLIENFCLFEMESHVTOAGVQWPNLGSLOLPPGLKRFCSLSPSS 340
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 SOSAGITGVSHHARLIENF--FETGTHSVTAAVQWYTTIGSLQRTPELKXSSHLILTSN 342
QY 341 WDYGHLPPHPAN-----FCIFIRGVSPLYSGWSQTPDLR 375
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 WDYRCTPPCPNLFYLFYFHRDEGSLCCPGXSXTPELK 382

RESULT 14
AAU31818
ID AAU31818 standard; protein; 382 AA.
AC AAU31818;
XX
XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #2309.
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX
XX Claim 20; Page 513; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
```

```
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 382 AA;
Query Match 45.6%; Score 927.5; DB 4; Length 382;
Best Local Similarity 55.8%; Pred. No. 1.1e-85;
Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;

QY 4 SLLPLRLCEGALSAHNRRLPGSSDSPASASPVAGITMCTHARLILYFVFLVEMFELHV 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4 SLLPLRLCEGALSAHNRRLPGSSDSPASASQVGTGKCHHTQLI-FVFLVEMGFPHI 62
QY 64 GQAGLELPTSDPSVSASQSYRTG--HHARLCLANFCGRNRVSLMCPWS----- 113
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 AQAGLELLTSDSPTL-ASQSAGI-TGVNHAWLFF--FCSRDTVSLCPGWSRVAXSRIT 118
QY 114 ----PELKQSTCLSLPKCWDYRRAV-PGLFILFPLRHRCPTLTQDEVQVCHSSLP-- 166
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 ATAPGLKXKFCFLSPSSRDYRHVPHPGNCIF-----GRDEVSPC-----WPGW 164
QY 167 -STPEIKHPPASQVAGTKDMHYTW----LIFIFIFNFRQSLNSVTQAGVQWRNLGS 221
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 FXTPLRLYPASASQSAEIIIGVSHHTWQEVFLFNLFYLRWLSDSVAQARVQRDLGS 224
QY 222 LQPLPPGKLFSCPSLLSSWDYRR--PPRLANFFVFLVEMGMTFARLILISGPCDL PASA 280
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 LQAPPFRKPFCLSLPSSWDYRRPPPPHPPANFFVFLVETGTVLARRVLISXPRDL PASA 284
QY 281 SOSAGITGVSHHARLIENFCLFEMESHVTOAGVQWPNLGSLOLPPGLKRFCSLSPSS 340
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 SOSAGITGVSHHARLIENF--FETGTHSVTAAVQWYTTIGSLQRTPELKXSSHLILTSN 342
QY 341 WDYGHLPPHPAN-----FCIFIRGVSPLYSGWSQTPDLR 375
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 WDYRCTPPCPNLFYLFYFHRDEGSLCCPGXSXTPELK 382

RESULT 15
AAU32707
ID AAU32707 standard; protein; 382 AA.
AC AAU32707;
XX
XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #3198.
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
```

PT vaccination, testing and therapy.

XX Claim 20; Page 652; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention

SQ Sequence 382 AA;

Query Match		45.6t;	Score 927.5;	DB 4;	Length 382;
Best Local Similarity		55.8t;	Pred. No. 1.1e-85;		
Matches 223;		Conservative 27;	Mismatches 101;	Indels 49;	Gaps 14;
Qy	4	SLLLPRLECGAISAHNRNLPGSSDSPASAPVAGITGCTHARLILYFPLVEMEFHV	63		
Db	4	SLLLPRLECGAIIAHNCLSGSSDSPASQVTGITGKCHTQLI-FVFLVEMGFHHI	62		
Qy	64	GQAGLELPTSDPSVSASQSARYTG--HHARLCLANFCGRNRVSLMCPWS	113		
Db	63	AQAGLELLTSDSPTL-ASQSAGI-TGVNHAWLFF--FCSRDTVSLCYPCGMSRVAXSRIT	118		
Qy	114	-----PELKQSTCLSLPKWDYRRAV-PGLFILFLRHRCPTLTQDEVQWCDHSSLOP--	166		
Db	119	ATSAPGLKXFCFSLPSSRDYRVPHPGNFCIF-----GRDEVSPC-----WPGW	164		
Qy	167	-STPEIKHPPASASQVAGTKDMHHTW---LIFIFLNFLOSLSNVTOAGVQWRNLGS	221		
Db	165	FXTPLRYPPASASQSAEIIQVSHHTWPQEVFLFNLFYLRWLSLSVAQARVORRDIGS	224		
Qy	222	LQPLPGFKLFCPSLLSSWDYRR-PPRLANFFVFLVEMGFTMPARLILISGPCDLPSA	280		
Db	225	LQAPPRFKPFCSLSPSSWDYRPPHPANFFVFLVETGTVLARRVLISXPRDLPSA	284		
Qy	281	SQSAGITGVSHHARLINFCLFEMESHVSVOAGVQWPNLGSLOPLPGLKRFPSCLSPSS	340		
Db	285	SQSAGITGVSHHTRLINF--FETGTHSVTAAVQWYITIGLSLOQRTPELKXSSHILTSN	342		
Qy	341	WDYGHLPHPAN-----FCIFIRGGVSPYLSGWSQTPDLR	375		
Db	343	WDYRCTPPCPNPLFIYLFYFHRDEGSLCCPGWSXTPELK	382		

Search completed: September 8, 2005, 10:37:53

Job time : 312 secs

THIS PAGE BLANK (USP 12)

Query Match	18.8%	Score 382.5;	DB 4;	Length 673;
Best Local Similarity	33.2%;	Pred. No. 6.4e-26;		
Matches	128;	Conservative	19;	Mismatches 111; Indels 127; Gaps 17;
QY	16	ISAHNRLRLPGSSDSDSPASASPVAGITGCMCTHARLILYFFLVEHFLVHGAGLELPTSD	75	
Db	371	VLAHCSLNLGGSDSPASVSRVAGITGMGRHMLI-YVFLTIQFHHVDQAGLKLLTSSD	429	
QY	76	PSVSASQSAARYTGTHARLCLANFCGRNRVSLMCPSPSPSPE-LKQSTCLSLPKCWDYRRAA	134	
Db	430	-----LPSWSPKVLGXQAXATPSPXXXF---V 453		
QY	135	VPGLFILFL-----RHRCP-----TITQDEVQVMDHSSLSQPSTPEIKHPSPASA	178	
Db	454	FGGFFFFFALFLRXALALTPLRECSGKFWLTAASTSW-QAILLPLSPVXLGLQAWA	512	
QY	179	SOVAGT-----KDMHHYTWLFIPIFNFLRQSLNSVTQAGVQWRNLGSLQPLPGCFKLPS	234	
Db	513	A-IPGXFWYFXRXHSFTMLI-----RLVSNXSPQV-----ICPPG-----	546	
QY	235	PSLLSSWDYRR-PPRLANFVFLVEMGFTFWARLILISGCDLPASASQAGITGVSHHA	293	
Db	547	--LPKWDDRRPPHPAXXXLFL-----GVFFP-----	572	
QY	294	RLIENFCLFMEHESHVTSQAGVQW---PNLASLQPLPPGLKRFKFSCLSPSSWDYGHLPHP	350	
Db	573	--FLLCFCXDRLLXHPG--WSAVASFGSLQPPQPGFKFKFSCLCLPCSWDYRHGPPFL	628	
QY	351	ANFCIFIRGGVSPYLSGWSQTFDLR 375		

Db 629 ANLCIFNRDVTSPCKSGWSQTPDLK 653

RESULT 2
A40201
artifact-warning sequence (translated ALU class C) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: C40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: C40201
A:Molecule type: DNA
A:Residues: 1-613 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potentially
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 613;
Best Local Similarity 33.4%; Pred. No. 5e-20;
Matches 98; Conservative 9; Mismatches 73; Indels 113; Gaps 8;

Qy 6 LLPLECNGAISAHNRNLRLPGSSDSPASAPVAGITGCTHARLILYFFLVEMFHLVQ 65
Db 436 LSPRLCNGAISAHCKLRLPGSRHSPASQVAGTTGARTTPG-XFFVFLVETGFHRSQ 494
Qy 66 AGLEPTSDDDPSVSASQSAARYRTGHHARLCLANFCGRNRVSLMCPSPWSPKQSTCLSLP 125
Db 495 DGLDLTLT-----SXSAR-----LCLP 510
Qy 126 KCWDYRRAAVPGULFILFLRHRCPRTLQDEVQWCHDSSLOPSTPEIKHPPASQVAGTK 185
Db 511 KCWDYRRETAPG-----
Qy 186 DMHHYTWLIFIF-----NFLRQSLNSVTQ-----AGVQWRNLGSLQPLDPGKLFSCPS 236
Db 523 ---XXXFLFFVFFRFDGVSLCQGWASVARSRLTASSASRVHAILLPQPPKXLGLOAPA 579
Qy 237 LLSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQASAGITGV 289
Db 580 L-----RPNFLYFXXRRGFTTVARVMSISXPRDPALASQASAGITGV 622

RESULT 4
E40201
artifact-warning sequence (translated ALU class E) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: E40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: E40201
A:Molecule type: DNA
A:Residues: 1-597 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potentially
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 14.3%; Score 291.5; DB 4; Length 597;
Best Local Similarity 30.5%; Pred. No. 6.6e-18;
Matches 118; Conservative 24; Mismatches 106; Indels 139; Gaps 20;

Qy 5 LLLPRLECNCAISAHNRNLRLPGSSD-SPASAPVAGITGCTHARLILYF-----FLVEM 58
Db 310 LLLPRPEYDGTISP--QTLPGFKQFSPVSASRVAGITGRHHAQLILYFXXRWGFSMLV 367
Qy 59 EFLHVGQAGLELPTSDDDPSVSASQSAARYT-----GHSYHTWXXXFFFEFCSCCGRSTMAQSHKLC 426
Db 368 KLVNSQPVIRPALNSQSAGIT-GMSYHTWXXXFFFEFCSCCGRSTMAQSHKLC 426
Qy 96 LANFCGRNRVSLMCPSPWSPKQSTCLSLPKWCDYRRAAVPGULFILFLRHRCPRTLQDE 155
Db 427 ---FLAGSSNLSLQ-----PEX-----LGLQAC-----ATPFSXYFI-----SRDG 461
Qy 156 V-----QWCDHSSLOPSTPEIKHPPASQVAGTKDMHHYTWLIFIFENFLRQSLNSVT 210
Db 462 VSPCWSWSQTPNLRXSAAPPXT-PKALGLQAXATP-----GXKXFFFLRRSFALVA 512
Qy 211 QAGVQWRNLGSLQPLPPGF-----KLFCSPCLSSWDYRR-PPRLANFFVFLVEMGFTMF 264
Db 513 QAGVRVHNLTA-----NFAWVQAAILSLCLSPSSWDYRHAPRPAN-FIFLWEMGFL-- 563
Qy 265 ARILISGCDLPASASQASAGITGVSHHARLIFNFCLEFEMESHVTOAGVQWPNLGSLOP 324
Db 564 -----HVGQAGLKLPTSGD--- 577

Db 629 ANLCIFNRDVTSPCKSGWSQTPDLK 653

RESULT 2
A40201
artifact-warning sequence (translated ALU class C) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: C40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: C40201
A:Molecule type: DNA
A:Residues: 1-613 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potentially
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 613;
Best Local Similarity 33.4%; Pred. No. 5e-20;
Matches 98; Conservative 9; Mismatches 73; Indels 113; Gaps 8;

Qy 6 LLPLECNGAISAHNRNLRLPGSSDSPASAPVAGITGCTHARLILYFFLVEMFHLVQ 65
Db 436 LSPRLCNGAISAHCKLRLPGSRHSPASQVAGTTGARTTPG-XFFVFLVETGFHRSQ 494
Qy 66 AGLEPTSDDDPSVSASQSAARYRTGHHARLCLANFCGRNRVSLMCPSPWSPKQSTCLSLP 125
Db 495 DGLDLTLT-----SXSAR-----LCLP 510
Qy 126 KCWDYRRAAVPGULFILFLRHRCPRTLQDEVQWCHDSSLOPSTPEIKHPPASQVAGTK 185
Db 511 KCWDYRRETAPG-----
Qy 186 DMHHYTWLIFIF-----NFLRQSLNSVTQ-----AGVQWRNLGSLQPLDPGKLFSCPS 236
Db 523 ---XXXFLFFVFFRFDGVSLCQGWASVARSRLTASSASRVHAILLPQPPKXLGLOAPA 579
Qy 237 LLSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQASAGITGV 289
Db 580 L-----RPNFLYFXXRRGFTTVARVMSISXPRDPALASQASAGITGV 622

RESULT 4
A40201
artifact-warning sequence (translated ALU class A) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: A40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: A40201
A:Molecule type: DNA
A:Residues: 1-627 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potentially
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 15.5%; Score 315.5; DB 4; Length 627;

```
Qy 325 LPPGLKRFSCLSLSPSSWDYGHLPHPHA 351
|||: |||||
Db 578 -PPRL-----LPKRWDRHELPHLA 597

RESULT 5
D40201
artifact-warning sequence (translated ALU class D) - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C;Accession: D40201
R;Claverie, J.M.
personal communication, 1992
A;Reference number: A40201
A;Accession: D40201
A;Molecule type: DNA
A;Residues: 1-579 <CLA>
R;Claverie, J.M.
Genomics 12, 838-841, 1992
A;Title: Identifying coding exons by similarity search: Alu-derived and other potential
A;Reference number: A40200; MUID:92241891; PMID:1572661
A;Contents: annotation
C;Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C;Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 9.7%; Score 197.5; DB 4; Length 579;
Best Local Similarity 28.1%; Pred. No. 1.4e-09;
Matches 81; Conservative 11; Mismatches 69; Indels 127; Gaps 10;
Qy 12 CN-GAISHARNRLRPGSSDSPASPVAGITGCMTHARLILYFFLVEMEFHVQAGLEL 70
|||: |||||
Db 404 CNHCSLS-----LDPPGVILLPQPFEXLGLQAYATRSXFLYLFVVEGFPHVAQAVLEL 459

Qy 71 PTSDPDSVSASQSARYRTGHHARICLANFCGRNVRVLMCPSPSPKQSTCLSLPKCWDY 130
|||: |||||
Db 460 -----LSWSTHLPW-----LLKCWDY 475

Qy 131 RRAAV-----PGLFILFHRCPILTQDEVQWCHSSLOPSTPEIKHPPASASQVAG 183
|||: |||||
Db 476 RRAAMLGXXXFRQGLGMLPRLEYRGAINA-----HXAL-----TSQV-- 512

Qy 184 TKDMHHVTWLIFFIFENFLRQSLNSVTQAGVQVWRNLGSLQPLPFGKLFSCPSLLSSWDY 243
|||: |||||
Db 513 -----XKSSCPSLLSSWDY 526

Qy 244 R-RPRLANFFV-FLVEMGFTMFARLILISGFCPLPASASQAGITGV 289
|||: |||||
Db 527 RLMLPDLANFCIFFLXRWGFMDLPRLLFLNSXAGAYTCHGSXSAGITGV 574

RESULT 6
S65657
alpha-1C-adrenergic receptor splice form 2 - human
N;Alternate names: alpha-1C-adrenoceptor isoform 2
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C;Accession: S65657; S65655
R;Tanaka, T.
submitted to the EMBL Data Library, July 1994
A;Reference number: S65656
A;Accession: S65657
A;Molecule type: mRNA
A;Residues: 1-499 <TAN>
A;Cross-references: UNIPROT:Q13675; EMBL:D32202; NID:G927208; PIDN:BAA06901.1; PID:G9272
R;Hirasawa, A.; Shibata, K.; Horie, K.; Takei, Y.; Obika, K.; Tanaka, T.; Muramoto, N.;
FEBS Lett. 363, 256-260, 1995
A;Title: Cloning, functional expression and tissue distribution of human alpha(1C)-adren
A;Reference number: S65654; MUID:95255557; PMID:7737411
A;Accession: S65655
A;Molecule type: mRNA
```

```
A;Residues: 424-499 <HIR>
A;Cross-references: EMBL:D32202
C;Genetics:
A;Gene: GDB:ADRA1C; ADRA1L1
A;Cross-references: GDB:128088; OMIM:104221
A;Map position: 8p21-8p11.2
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; neurotransmitter receptor

Query Match 8.6%; Score 175; DB 2; Length 499;
Best Local Similarity 58.0%; Pred. No. 1.1e-07;
Matches 40; Conservative 5; Mismatches 20; Indels 4; Gaps 1;
Qy 9 RLECNGAISHARNRLRPGSSDSPASPVAGITGMC-----THARLILYFFLVEMEFHVHG 64
|||: |||||
Db 431 RLECSGMILAHNCNLRPLGSRDSPASASQAAGTTGDPVPPGRHQALIFVLVETGPFHVHG 490

Qy 65 QAGLELPTS 73
|||: |||||
Db 491 QDDLDDLLTS 499

RESULT 7
I54375
gene NF2 protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 16-Aug-2004
C;Accession: I54375
R;Arakawa, H.; Hayaashi, N.; Nagase, H.; Ogawa, M.; Nakamura, Y.
Hum. Mol. Genet. 3, 565-568, 1994
A;Title: Alternative splicing of the NF2 gene and its mutation analysis of breast and co.
A;Reference number: I54375; MUID:94348501; PMID:8069299
A;Accession: I54375
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-46 <RES>
A;Cross-references: UNIPROT:Q16230; GB:S73853; NID:G688372; PIDN:AAB31736.1; PID:G688373
C;Genetics:
A;Gene: GDB:NF2
A;Cross-references: GDB:I20232; OMIM:101000
A;Map position: 22q12.2-22q12.2
C;Superfamily: laminin-type EGF-like homology

Query Match 8.2%; Score 167; DB 2; Length 46;
Best Local Similarity 76.2%; Pred. No. 3.6e-08;
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
Qy 302 FEMESHVTOAGVQVWNLGSLQPLPGLKRFSCLSLSPSSWDY 343
|||: |||||
Db 4 FNCESCSVTLAGVQWRDLGLLQPLPPKFKRFSCLSLSPSSWDY 45

RESULT 8
I78885
serine/threonine-specific protein kinase (EC 2.7.1.1-) STK2 - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: I78885
R;Levedakou, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Welcsh, P.L.; Simmo
Oncogene 9, 1977-1988, 1994
A;Title: Two novel human serine/threonine kinases with homologies to the cell cycle regu
A;Reference number: I58396; MUID:94268838; PMID:8208544
A;Accession: I78885
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-841 <RES>
A;Cross-references: UNIPROT:P51957; GB:L20321; NID:G348244; PIDN:AAA36658.1; PID:G348245
C;Genetics:
A;Gene: GDB:STK2
A;Cross-references: GDB:374125
A;Map position: 3p21.1-3p21.1
C;Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homol
C;Keywords: phosphotransferase
```

```
Query Match      7.4%; Score 151.5; DB 4; Length 301;  
Best Local Similarity 22.4%; Pred. No. 7.6e-06;  
Matches        66; Conservative    15; Mismatches   45; Indels    169; Gaps     10;
```

QY 6 LLPRLECNCAISAHNRNLPGSSDPSASPFVAGITGMCT-----HARLLIYPFLVEMEF 60
::|||:| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

Best Local Similarity 49.3%; Pred. NO. 2.8e-05;
Matches 37; Conservative 8; Mismatches 25; Indels 5; Gaps 2;

QY 6 LPLRCEGAI SAHRLNLPGGSDSPASAPVAGTCMTHARLILYFLVEMEFHVQG 65
| | | : | | | | | | | | | | : |
Db 5 LMPFLEYGTISAHCNLRLPGSSDRASASRAAGITGV-SHCARPOMLEDPDFDLL---- 59


```
Db 152 -----STAVEFCKKSC-PNPGEIRN---GQIDVPG-----G 179
QY 193 LIFIFIENF-----LRQSLNS---VTOAGVQWENLGSLOPLPEGFKLFSCPSLLSSWD 242
Db 180 ILFGATISFCNTGKLFGSTSSFCLSGSSVQWSD-----PLPECREIY-CPA----- 227
QY 243 YRRPPLANFFV-----FLVEMGFTMFARLILI-----SGP----- 273
Db 228 ---PPQIDNGIIQGERDHVGRQSVTYACNKGFTMIGEHSIYCTVNNDGEGWSGPPPECR 284
QY 274 -----CDLP-----ASASQAGITGVSHH----- 292
Db 285 GKSLTSKVPPTVQKPTTVNVPTTEVSP7SQKTTTKTTTFNAQATRSTPEVSR7TKHFETT 344
QY 293 -----ARLIFNCLFEMESHVTOAGVQWPNLGSLOPLPGLKRFSCLSLPSW 341
Db 345 PNKSGTTSGTTRLSG-----SRPVTQAGMRWCDRSSLQSRTPGPKRSFHFSLPSSW 397
QY 342 DY 343
Db 398 YV 399
```

Search completed: September 8, 2005, 10:45:33
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 06:54:01 ; Search time 376 Seconds
(without alignments)
510.717 Million cell updates/sec

Title: US-10-092-934-10
Perfect score: 2034
Sequence: 1 MEFSLLPLRLCNGAISAHRR.....FIRGGVSPYLXSGWSQTFDLR 375

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	2	O60448
2	595	29.3	587	1	ALUJ2_HUMAN
3	568	27.9	587	1	ALUJ3_HUMAN
4	549.5	27.0	591	1	ALUJ8_HUMAN
5	498.5	24.5	591	1	ALUJ1_HUMAN
6	389	19.1	585	1	ALUJ5_HUMAN
7	377.5	18.6	158	2	Q8NAL9
8	357.5	17.6	593	1	ALUJ6_HUMAN
9	356.5	17.5	603	1	ALUJ4_HUMAN
10	349	17.2	157	2	Q8N845
11	337.5	16.6	164	2	Q6ZTS0
12	332	16.3	176	2	Q6ZRN2
13	331	16.3	170	2	Q6ZV14
14	330.5	16.2	593	1	ALUJ7_HUMAN
15	325	16.0	143	2	Q6ZSJ7
16	320	15.7	195	2	Q6ZUM2
17	316	15.5	152	2	Q8NX85
18	303.5	14.9	239	2	Q6ZMA9
19	303	14.9	124	2	Q6ZVF6
20	301	14.8	156	2	Q8N266
21	295	14.5	193	2	Q6ZVX8
22	294.5	14.5	148	2	Q6ZUC5
23	290	14.3	123	2	Q6ZV58
24	287.5	14.1	166	2	Q6ZUC0
25	287	14.1	180	2	Q8N7M7
26	286.5	14.1	130	2	Q6ZSV2
27	286.5	14.1	131	2	Q6ZTD3
28	283	13.9	125	2	Q6ZTE1
29	281.5	13.8	147	2	Q6ZUA2
30	280	13.8	165	2	Q6ZTF6
31	279.5	13.7	132	2	Q6ZS28

ALIGNMENTS

RESULT 1

O60448	PRELIMINARY;	PRT;	375 AA.
AC	O60448;		
DT	01-AUG-1998 (TrEMBLrel. 07, Created)		
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Neuronal thread protein AD7c-NTP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Neuronal;		
RX	MEDLINE=98064067; PubMed=9399956;		
RA	Monte S.M., Ghanbari K., Frey W.H., Beheshti I., Averbach P.,		
RA	Hauser S.L., Ghanbari H.A., Wands J.R.;		
RT	"Characterization of the AD7C-NTP cDNA expression in Alzheimer's		
RT	disease and measurement of a 41-kD protein in cerebrospinal fluid."		
RL	J. Clin. Invest. 100:3093-3104(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Neuronal;		
RA	de la Monte S.M., Wands J.R.;		
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF010144; AAC08737.1;		
DR	GO; GO:0005615; C:extracellular space; TAS.		
DR	GO; GO:0016021; C:integral to membrane; TAS.		
DR	GO; GO:0006915; P:apoptosis; TAS.		
DR	GO; GO:0007417; P:central nervous system development; TAS.		
SQ	SEQUENCE 375 AA; 41720 MW; 955443950ASBFEF CRC64;		
Query Match 100.0%; Score 2034; DB 2; Length 375;			
Best Local Similarity 100.0%; Pred. NO. 1.9e-171;			
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MEFSLLPLRLCNGAISAHRRALPGSSDSPASPSVAGITGMCTHARLLYFLVEMEF	60
Db	1	MEFSLLPLRLCNGAISAHRRALPGSSDSPASPSVAGITGMCTHARLLYFLVEMEF	60
Qy	61	LHVQAGLEIPTSDDPSVSASQARVYTGHHARLLCLANFCGRNRVSLMCPSPKQST	120
Db	61	LHVQAGLEIPTSDDPSVSASQARVYTGHHARLLCLANFCGRNRVSLMCPSPKQST	120
Qy	121	CLSIPKCDWYRRAAVPGLFILFLRHRCPTLTQDEVQCDHSSLPQSTPIKHPASASQ	180
Db	121	CLSIPKCDWYRRAAVPGLFILFLRHRCPTLTQDEVQCDHSSLPQSTPIKHPASASQ	180
Qy	181	VAGTKDMHYTWLIFIFINFLQSLNSVTQAGVQWENLQSLQPLPGFKLFSCPSLSS	240
Db	181	VAGTKDMHYTWLIFIFINFLQSLNSVTQAGVQWENLQSLQPLPGFKLFSCPSLSS	240
Qy	241	WDYRRPRLANFFVFLVEMGFTWPARLLISGPCDLPASASQSGITGVSHHARLLIFNFC	300

Q6zwd5 homo sapien
Q6zwd4 homo sapien
Q8n8c2 homo sapien
Q6zuko0 homo sapien
Q9h387 homo sapien
Q6zul3 homo sapien
Q6zge8 homo sapien
Q6z897 homo sapien
Q6zyu3 homo sapien
Q6zuw8 homo sapien
Q6zqr8 homo sapien
Q6z553 homo sapien
Q6zt71 homo sapien
Q6zsu7 homo sapien

Db 241 WYRRPRLANFFVLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIENFC 300
 Qy 301 LFEMESHVSQTQAGVQWPNLGSQPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
 Db 301 LFEMESHVSQTQAGVQWPNLGSQPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
 Qy 361 VSPYLSGWSQTPDLR 375
 Db 361 VSPYLSGWSQTPDLR 375

RESULT 2

ALU2 HUMAN
 ID ALU2 HUMAN STANDARD; PRT; 587 AA.
 AC P39189;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SB sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;
 RT "Alu alert."
 RL Nature 371:752-752(1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 potentially misleading protein sequences.";
 RL Genomics 12:838-841(1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely
 connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202(1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121(1991).
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
 into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
 consensus sequences have been constituted that contain all six
 frames conceptual translations of each of these classes of Alu
 repeats.
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
 codon, 'XXX' is used to separate the various translation phases.
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further
 pollution of protein sequence databases with Alu-derived amino
 acid sequences.
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and
 primate genomes with an average spacing of 4 kb. Some of them are
 actively transcribed by pol III. Normal transcripts may contain
 Alu-derived sequences in 5' or 3' untranslated regions. However,
 cDNA libraries also contain partial and/or rearranged cDNAs
 ligated with Alu-derived sequence in any orientation. Although Alu
 elements (especially situated on the complementary strand) have a
 great potential to create additional/alternative exons,
 consideration should be given to the possibility that the presence
 of an Alu in an open reading frame may have resulted from a
 cloning artifact or may be due to misinterpretation of sequencing
 data. This point has been overlooked on several occasions, with
 the consequence of erroneous Alu-derived amino acid sequences

CC being reported.
 CC -!- CAUTION: Any significant similarity of a putative protein sequence
 with an Alu-translated entry must be taken as a warning that a
 part of Alu repeat may have been artifactually included in the
 coding nucleotide sequence.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; U14568; -; NOT ANNOTATED_CDS.
 DR Hypothetical protein.
 KW DOMAIN 1 96 Frame-1.
 FT DOMAIN 100 194 Frame-2.
 FT DOMAIN 198 292 Frame-3.
 FT DOMAIN 296 391 Frame-4.
 FT DOMAIN 395 489 Frame-5.
 FT DOMAIN 493 587 Frame-6.
 SQ SEQUENCE 587 AA; 63703 MW; 3EAB3E3E3929203 CRC64;
 Query Match 29.3%; Score 595; DB 1; Length 587;
 Best Local Similarity 41.5%; Pred. No. 5.7e-44;
 Matches 153; Conservative 14; Mismatches 96; Indels 106; Gaps 8;
 Qy 6 LLPRLECNAGISAHRNLRPLFGSSDPSASAPVAGITGCTHARILYFFLVEMEFHVGQ 65
 Db 304 LSPRECGSAISAHCKLRPLGSRHSPASASRVAGTGTGARHARLI-FVFLVETGHRVSQ 362
 Qy 66 AGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNYSVLMCPSPKQSTCLSLP 125
 Db 363 DGLDLLT-----SXSAR-----LGLP 378
 Qy 126 KCWDYRRAAVPGLFILFRLRHCPTLTQDEVQVQCDHSSLOPSTPEIKHPASASQVAGTK 185
 Db 379 KCWDYRREPPRPAXXFF-----XDGVSIC-----RPGWSAVARSRLTASAS--- 421
 Qy 186 DMHHYTWLIFIFINFLRSLNSVTQAGVQWNLGSLQPLPPGKFLPCFSLSSWDYRR 245
 Db 422 -----RVHAILLPQPEXGLQADATP----- 444
 Qy 246 PRLANFFVLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIENFCLEME 305
 Db 445 -----GXFLYFXRRRGFTVTLARMYSISXPRDPPASASQAGITGVSHHARXXX---FFETE 497
 Qy 306 SHSVTQAGVQWPNLGSQPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGVSPYL 365
 Db 498 SRVAQAGVQWRDLGSLQAPPGFTFPFSCLSLPSWDYRRPPRPPANFCIFSRDGVSPCX 557
 Qy 366 SGWSQTPDL 374
 Db 558 PGWSRSPDL 566
 RESULT 3
 ALU3 HUMAN
 ID ALU3 HUMAN STANDARD; PRT; 587 AA.
 AC P39190;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alu subfamily SB1 sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;

RT "Alu alert."; 6
 RL Nature 371:752-752(1994). 304
 RP CONCEPT. 66
 RA MEDLINE=92241891; PubMed=1572661; 363
 RX Claverie J.-M.; 126
 RT "Identifying coding exons by similarity search: alu-derived and other 379
 RT potentially misleading protein sequences."; 186
 RL Genomics 12:838-841(1992). 415
 RN [3] 246
 RP ALU FAMILIES CLASSIFICATION. 445
 RX MEDLINE=88333009; PubMed=3138422; 306
 RA Quentín Y.; 498
 RT "The Alu family developed through successive waves of fixation closely 366
 RT connected with primate lineage history."; 558
 RL J. Mol. Evol. 27:194-202(1988). 557
 RN [4] 365
 RP ALU FAMILIES CLASSIFICATION. 365
 RX MEDLINE=91178815; PubMed=1706781; 365
 RA Jurka J., Milosavljevic A.; 365
 RT "Reconstruction and analysis of human Alu genes."; 365
 RL J. Mol. Evol. 32:105-121(1991). 365
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall 365
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning 365
 CC consensus sequences have been constituted that contain all six 365
 CC frames conceptual translations of each of these classes of Alu 365
 CC repeats. 365
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop 365
 CC codon, 'XXX' is used to separate the various translation phases. 365
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further 365
 CC pollution of protein sequence databases with Alu-derived amino 365
 CC acid sequences. 365
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and 365
 CC primate genomes with an average spacing of 4 kb. Some of them are 365
 CC actively transcribed by pol III. Normal transcripts may contain 365
 CC Alu-derived sequences in 5' or 3' untranslated regions. However, 365
 CC cDNA libraries also contain partial and/or rearranged cDNAs 365
 CC ligated with Alu-derived sequence in any orientation. Although Alu 365
 CC elements (especially situated on the complementary strand) have a 365
 CC great potential to create additional/alternative exons, 365
 CC consideration should be given to the possibility that the presence 365
 CC of an Alu in an open reading frame may have resulted from a 365
 CC cloning artifact or may be due to misinterpretation of sequencing 365
 CC data. This point has been overlooked on several occasions, with 365
 CC the consequence of erroneous Alu-derived amino acid sequences 365
 CC being reported. 365
 CC -!- CAUTION: Any significant similarity of a putative protein sequence 365
 CC with an Alu-translated entry must be taken as a warning that a 365
 CC part of Alu repeat may have been artifactually included in the 365
 CC coding nucleotide sequence. 365
 CC ----- 365
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration 365
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation - 365
 CC the European Bioinformatics Institute. There are no restrictions on its 365
 CC use by non-profit institutions as long as its content is in no way 365
 CC modified and this statement is not removed. Usage by and for commercial 365
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/> 365
 CC or send an email to license@ebi.ac.uk). 365
 CC ----- 365
 CC EMBL; U14569; -; NOT_ANNOTATED_CDS. 365
 DR Hypothetical protein. 365
 KW DOMAIN 1 96 Frame-1. 365
 FT DOMAIN 100 194 Frame-2. 365
 FT DOMAIN 198 292 Frame-3. 365
 FT DOMAIN 296 391 Frame-4. 365
 FT DOMAIN 395 489 Frame-5. 365
 FT DOMAIN 493 587 Frame-6. 365
 SQ SEQUENCE 587 AA; 63573 MW; 85C4155726DEF235 CRC64; 365
 Query Match 27.9%; Score 568; DB 1; Length 587;
 Best Local Similarity 41.5%; Pred. No. 1.4e-41;
 Matches 153; Conservative 15; Mismatches 95; Indels 106; Gaps 10;

QY 6 LLPRLECNAGISAHRNLRUPGSSDSDSPASASPVAGITGCTHARILLYFFLVMEFLHVQ 65
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 304
 LSPLRECSGGISAHCKRLPGRSRHSPASASQVAGTTGCAHYARLI-FVFLVETGCFHRSR 362
 QY 66 AGLELPTSDPPSVASASQARYRTGHARLCLANFCGNGRNVRLMCPSPSPKQSTCLSLP 125
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 363
 DGJDLIT-----SASAR-----LGLP 378
 QY 126 KWDYRRRAAVPGLFILFLHRCPTLTQDEVOMCDHSSLOPSTPEIKHPPASASQVAGTK 185
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 379
 KWDYRREP RPAXXFF-----XDGVSIC-----RPGWSA--VAGSR 414
 QY 186 DMHYTWLIPFIFNFIROSLNSVTOAGVOMRNLGSLQPLPPGFKLFSCPSLSSWDYRR 245
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 415
 -----LTASSASRVHA-----ILLPQPKXGLQAPATTP----- 444
 QY 246 PPRLANFFVFLVEMGTMTFARLILISGCDLPASASQAGITGVSHHARLIFNFCLEME 305
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 445
 -----GXFLYFXRXRGFTVLAGMVSISXPRDPPASASQAGITGVSHHARXXX---PFETE 497
 QY 306 SHSVTOAGVOMRNLGSLQPLPPGFKLFSCPSLSSWDYHGLPHHPANFCIFIRGGVSPYL 365
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 498
 SRVAQGVOMRDLGSLQAPPPGFTFSCSLSSWDYRRPPLRPANFCIFSRDGVSPFX 557
 QY 366 SCMSQTPDL 374
 DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 558
 PGWSRSPDL 566
 RESULT 4
 ID ALU8 HUMAN STANDARD; PRT; 591 AA.
 AC P39195;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SX sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021759; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;
 RL "Alu alert."; RT
 RL Nature 371:752-752(1994). [2]
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other
 RT potentially misleading protein sequences.";
 RL Genomics 12:838-841(1992). [3]
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=88333009; PubMed=3138422;
 RA Quentín Y.;
 RT "The Alu family developed through successive waves of fixation closely
 RT connected with primate lineage history.";
 RL J. Mol. Evol. 27:194-202(1988). [4]
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes.";
 RL J. Mol. Evol. 32:105-121(1991).
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
 CC consensus sequences have been constituted that contain all six
 CC frames conceptual translations of each of these classes of Alu
 CC repeats.
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
 CC codon, 'XXX' is used to separate the various translation phases.
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further
 CC pollution of protein sequence databases with Alu-derived amino
 CC acid sequences.
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and
 CC primate genomes with an average spacing of 4 kb. Some of them are
 CC actively transcribed by pol III. Normal transcripts may contain
 CC Alu-derived sequences in 5' or 3' untranslated regions. However,
 CC cDNA libraries also contain partial and/or rearranged cDNAs
 CC ligated with Alu-derived sequence in any orientation. Although Alu
 CC elements (especially situated on the complementary strand) have a
 CC great potential to create additional/alternative exons,
 CC consideration should be given to the possibility that the presence
 CC of an Alu in an open reading frame may have resulted from a
 CC cloning artifact or may be due to misinterpretation of sequencing
 CC data. This point has been overlooked on several occasions, with
 CC the consequence of erroneous Alu-derived amino acid sequences
 CC being reported.
 CC -!- CAUTION: Any significant similarity of a putative protein sequence
 CC with an Alu-translated entry must be taken as a warning that a
 CC part of Alu repeat may have been artifactually included in the
 CC coding nucleotide sequence.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; U14569; -; NOT_ANNOTATED_CDS.
 DR Hypothetical protein.
 KW DOMAIN 1 96 Frame-1.
 FT DOMAIN 100 194 Frame-2.
 FT DOMAIN 198 292 Frame-3.
 FT DOMAIN 296 391 Frame-4.
 FT DOMAIN 395 489 Frame-5.
 FT DOMAIN 493 587 Frame-6.
 SQ SEQUENCE 587 AA; 63573 MW; 85C4155726DEF235 CRC64;

```

CC      codon. 'XXX' is used to separate the various translation phases.
CC      -!- CAUTION: This Alu entry is provided in order to avoid the further
CC      pollution of protein sequence databases with Alu-derived amino
CC      acid sequences.
CC      -!- CAUTION: Alu repetitive sequences are interspersed in human and
CC      primate genomes with an average spacing of 4 kb. Some of them are
CC      actively transcribed by pol III. Normal transcripts may contain
CC      Alu-derived sequences in 5' or 3' untranslated regions. However,
CC      cDNA libraries also contain partial and/or rearranged cDNAs
CC      ligated with Alu-derived sequence in any orientation. Although Alu
CC      elements (especially situated on the complementary strand) have a
CC      great potential to create additional/alternative exons,
CC      consideration should be given to the possibility that the presence
CC      of an Alu in an open reading frame may have resulted from a
CC      cloning artifact or may be due to misinterpretation of sequencing
CC      data. This point has been overlooked on several occasions, with
CC      the consequence of erroneous Alu-derived amino acid sequences
CC      being reported.
CC      -!- CAUTION: Any significant similarity of a putative protein sequence
CC      with an Alu-translated entry must be taken as a warning that a
CC      part of Alu repeat may have been artifactually included in the
CC      coding nucleotide sequence.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U14574; -; NOT ANNOTATED_CDS.
CC      KW      Hypothetical protein.
CC      FT      DOMAIN 1 96 Frame-1.
CC      FT      DOMAIN 100 195 Frame-2.
CC      FT      DOMAIN 199 294 Frame-3.
CC      FT      DOMAIN 298 393 Frame-4.
CC      FT      DOMAIN 397 492 Frame-5.
CC      FT      DOMAIN 496 591 Frame-6.
CC      SQ      SEQUENCE 591 AA; 64395 MW; AC8154ADBA6BB280 CRC64;
CC
CC      Query Match      27.0%; Score 549.5; DB 1; Length 591;
CC      Best Local Similarity 38.9%; Pred. No. 6.1e-40;
CC      Matches 150; Conservative 15; Mismatches 84; Indels 137; Gaps 11;
CC
CC      QY      6 LLPLRECGAISAHNRLPGSSDSPASASPVAGITGCTHARLILYFFFLVEMERFLHVQ 65
CC      DB      306 LSPRLCEGSAISAHNRLPGSSDSPASASPVAGITGARHARLI-FVFLVETGFHHVQ 364
CC
CC      QY      66 AGLELPTSD--PSVSASQSARYRTGTHARLCLANFCGRNVRSLMCPSSW-----P 114
CC      DB      365 AGLELTSGLDLPKVLGLQAXATPG---LRPXXXFFXGDLGSLCRPGMSAVARSRLTAT 420
CC      QY      115 ELKQSTCLSLPKCWD---YRAAVPGILFLLFRLHRCPTLTQDEVQWCHSSLPSTPE 170
CC      DB      421 SASRVAQILLQPPPELXGLQARATPGXFLYFXRRGRFTMLAR-----LVNSNXPQ 471
CC      QY      171 IKHPPASASQVAGTKDMHHYTLWLIFFLFLRQSLNSVTQAGVQWRNLGSLQLPPGFK 230
CC      DB      472 V-----IF----- 474
CC      QY      231 LFSCLPSSLLSSWDYRR-PPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGV 289
CC      DB      475 -----LPKCWDYRREPR-----PASARXXX----- 495
CC      QY      290 SHHARLIENFLCFEMESHVSUQAGVQWRNLGSLQLPPLGKRFKFCLSLPSSWDYGHLPDPH 349
CC      DB      496 -----FFETESRSVAQGVQWRDLGSLQPPPPGKRFKFCLSLPSSWDYRRAPPR 544
CC      QY      350 PANFCIFIRGGVSPYLSGMSQTPDLR 375
CC      DB      545 PANFCIFRSDGVSPCWPQWSRTPDLR 570

```

```

RESULT 5
ALU1_HUMAN
ID      ALU1_HUMAN      STANDARD;      PRT;      591 AA.
AC      P39188;
DT      01-FEB-1995 (Rel. 31, Created)
DT      01-FEB-1995 (Rel. 31, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Alu subfamily J sequence contamination warning entry.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
SEQUENCE FROM N.A.
RP      MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
RA      Claverie J.-M., Makalowski W.;
RT      "Alu alert.";
RL      Nature 371:752-752 (1994).
RN      [2]
CONCEPT.
RX      MEDLINE=92241891; PubMed=1572661;
RA      Claverie J.-M.;
RT      "Identifying coding exons by similarity search: alu-derived and other
RT      potentially misleading protein sequences.";
RL      Genomics 12:838-841 (1992).
RN      [3]
ALU FAMILIES CLASSIFICATION.
RX      MEDLINE=88333009; PubMed=3138422;
RA      Quentin Y.;
RT      "The Alu family developed through successive waves of fixation closely
RT      connected with primate lineage history.";
RL      J. Mol. Evol. 27:194-202 (1988).
RN      [4]
ALU FAMILIES CLASSIFICATION.
RX      MEDLINE=91178815; PubMed=1706781;
RA      Jurka J., Milosavljevic A.;
RT      "Reconstruction and analysis of human Alu genes.";
RL      J. Mol. Evol. 32:105-121 (1991).
CC      -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
CC      into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
CC      consensus sequences have been constituted that contain all six
CC      frames conceptual translations of each of these classes of Alu
CC      repeats.
CC      -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
CC      codon. 'XXX' is used to separate the various translation phases.
CC      -!- CAUTION: This Alu entry is provided in order to avoid the further
CC      pollution of protein sequence databases with Alu-derived amino
CC      acid sequences.
CC      -!- CAUTION: Alu repetitive sequences are interspersed in human and
CC      primate genomes with an average spacing of 4 kb. Some of them are
CC      actively transcribed by pol III. Normal transcripts may contain
CC      Alu-derived sequences in 5' or 3' untranslated regions. However,
CC      cDNA libraries also contain partial and/or rearranged cDNAs
CC      ligated with Alu-derived sequence in any orientation. Although Alu
CC      elements (especially situated on the complementary strand) have a
CC      great potential to create additional/alternative exons,
CC      consideration should be given to the possibility that the presence
CC      of an Alu in an open reading frame may have resulted from a
CC      cloning artifact or may be due to misinterpretation of sequencing
CC      data. This point has been overlooked on several occasions, with
CC      the consequence of erroneous Alu-derived amino acid sequences
CC      being reported.
CC      -!- CAUTION: Any significant similarity of a putative protein sequence
CC      with an Alu-translated entry must be taken as a warning that a
CC      part of Alu repeat may have been artifactually included in the
CC      coding nucleotide sequence.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U14567; -; NOT ANNOTATED_CDS.

CC DR Hypothetical protein_96 Frame-1. 100 195 Frame-2. 100 195 Frame-3. 199 294 Frame-4. 298 393 Frame-5. 397 492 Frame-6. 496 591 Frame-6. 591 63790 MW; 665D395735519D95 CRC64;

CC Query Match 24.5%; Score 498.5; DB 1; Length 591; Best Local Similarity 38.6%; Pred. No. 2e-35; Matches 144; Conservative 15; Mismatches 103; Indels 111; Gaps 11;

Qy 6 LLPRLENGAISAHRNLRPGSSDPASAPVAGITGCTHARLILYFFLVEMFELVHGQ 65
Db 306 LSPRLCSGATTAHCSLDLPGSSDPASAPVAGITGCTHARLILYFFLVEMFELVHGQ 364

Qy 66 AGLELPTSDPSVASQASARYRTG--HHARLCLANFCGRNRSVLMCPSPKQSTCLS 123
Db 365 AGLELGSDDPPASQASAGI-TGVSHR--XXXFFXDRVSLCRPGNSA-----411

Qy 124 LPKCWYRRAAVPGLFLFLFRLHRCPTLTQDEVQWCDHSSLOPSTPPIKHPASASQVAG 183
Db 412 -----VARGSLTAASSTRAQ-----AIIIPQPE-----435

Qy 184 TKDMHVTWLIIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFPCPSLLSNDY 243
Db 436 -----XLGLQAR-----ATTFG-----447

Qy 244 RRPRLANFFVFLVEMGFTMFARLILIS-GPCDLPASASQASAGITGVSHHARLIFNFCFLF 302
Db 448 -----XFLYFLXRRGFTMLPRVNSMAQVILPPRPK---VLQQAATAPGXKXFF 497

Qy 303 EMBSHSVTQAGVQWRNLGSLQPLPPGKLFPCPSLLSNDYHLPHPANFCIFIRGVS 362
Db 498 ETGSRVAQAGVQWRNLGSLQPLPPGKLFPCPSLLSNDYHLPHPANFCIFIRGVS 557

Qy 363 PYLSGWSQTPDLR 375
Db 558 LCCPGWSRTPELK 570

RESULT 6

ALU5 HUMAN STANDARD; PRT; 585 AA.
AC P39192;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE ALU subfamily SC sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
RA Claverie J.-M., Makalowski W.;
RT "Alu alert."
RL Nature 371:752-752(1994).
RN [2]
RP CONCEPT.
RX MEDLINE=92241891; PubMed=1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other potentially misleading protein sequences."
RL Genomics 12:838-841(1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.

RX MEDLINE=88333009; PubMed=3138422;
RA Quentin Y.;
RT "The Alu family developed through successive waves of fixation closely connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).

RN [4] ALU FAMILIES CLASSIFICATION.
RP MEDLINE=91178815; PubMed=1706781;
RX Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).

CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning consensus sequences have been constituted that contain all six frames conceptual translations of each of these classes of Alu repeats.

CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop codon, 'XX' is used to separate the various translation phases.
CC -!- CAUTION: This Alu entry is provided in order to avoid the further pollution of protein sequence databases with Alu-derived amino acid sequences.

CC -!- CAUTION: Alu repetitive sequences are interspersed in human and primate genomes with an average spacing of 4 kb. Some of them are actively transcribed by pol III. Normal transcripts may contain Alu-derived sequences in 5' or 3' untranslated regions. However, cDNA libraries also contain partial and/or rearranged cDNAs ligated with Alu-derived sequence in any orientation. Although Alu elements (especially situated on the complementary strand) have a great potential to create additional/alternative exons, consideration should be given to the possibility that the presence of an Alu in an open reading frame may have resulted from a cloning artifact or may be due to misinterpretation of sequencing data. This point has been overlooked on several occasions, with the consequence of erroneous Alu-derived amino acid sequences being reported.

CC -!- CAUTION: Any significant similarity of a putative protein sequence with an Alu-translated entry must be taken as a warning that a part of Alu repeat may have been artifactually included in the coding nucleotide sequence.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U14571; -; NOT ANNOTATED_CDS.
KW Hypothetical protein.
FT DOMAIN 1 95 Frame-1.
FT DOMAIN 99 193 Frame-2.
FT DOMAIN 197 291 Frame-3.
FT DOMAIN 295 389 Frame-4.
FT DOMAIN 393 487 Frame-5.
FT DOMAIN 491 585 Frame-6.
SQ SEQUENCE 585 AA; 63957 MW; 46EB8C4F493650A7 CRC64;

Query Match 19.1%; Score 389; DB 1; Length 585;
Best Local Similarity 33.9%; Pred. No. 9.5e-26;
Matches 130; Conservative 23; Mismatches 100; Indels 130; Gaps 17;

Qy 17 SAHRNLRPGSSDPASAPVAGITGCTHARLILYFFLVEMFELVHGQAGLELPTSDDP 76
Db 189 SVSKXXXPGVAHACNPSITLGGGGRITRSR-----DRDHPGQHG-----ETP 232

Qy 77 SVSASQASARYTGHARLCLANFCGR-----NRV-----SLMC-PSWSPKQST 120
Db 233 SLTKIKLAGRG--ARL-XSQLGLRLQENRLNPGGGCGSEPRSRHCTPAW---RQSE 285

Qy 121 CLSLPKCWYRRAAVPGLFLFLFRLHRCPTLTQDEVQWCDHSSLOPSTPPIKHPASASQ 180
Db 286 TPSQKK-----XXXFFLRSL-----ALSPGW-----SA 309

being reported.
 -!- CAUTION: Any significant similarity of a putative protein sequence with an Alu-translated entry must be taken as a warning that a part of Alu repeat may have been artifactually included in the coding nucleotide sequence.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; U14572; -; NOT_ANNOTATED_CDS.
 ProDom; PD003738; GIDA; 1.
 Hypothetical protein.
 DOMAIN 1 97 Frame-1.
 DOMAIN 101 196 Frame-2.
 DOMAIN 200 295 Frame-3.
 DOMAIN 299 395 Frame-4.
 DOMAIN 399 494 Frame-5.
 DOMAIN 498 593 Frame-6.
 SEQUENCE 593 AA; 64603 MW; 136EF344AACD12A2 CRC64;
 Query Match 17.6%; Score 357.5; DB 1; Length 593;
 Best Local Similarity 48.2%; Pred. No. 5.9e-23;
 Matches 95; Conservative 6; Mismatches 47; Indels 49; Gaps 8;
 ;
 QY 4 SLLPRLCEGCAISAHNRLPGSSDSPASPVAGITGCTHARLLYFLVEMELHV 63
 D 405 SLLPRLCEGCAISAHNRLPGSSDSPASASRVRAGITGMRHARLLI-FVFLVETGFLHV 463
 QY 64 GQAGLEPTSDPSVASQSARYRTG--HHARLCLANFCGRRNRLVSLMCPGWS----- 113
 D 464 GQAGLEPTSDPSVASQSAGI-TGVSHRAR-----XXXFFETEFSCPGWSAMARSLT 519
 QY 114 -----PELKOSTCLSLPKCWDYRRAAVPGILFLFLRHRCPTLLPQDEVQW 158
 D 520 ATSASRVQAILLPQPEX-----LGLQAC-----ATTPGXFLYEXRRGFSMLVR----- 564
 QY 159 CDHSSLQSPSTEIKHPP 175
 D 565 ----LVNSNRPOVIRPP 577
 RESULT 9
 ALU4_HUMAN
 ID ALU4_HUMAN STANDARD; PRT; 603 AA.
 AC P39191;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Alu subfamily SB2 sequence contamination warning entry.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
 RA Claverie J.-M., Makalowski W.;
 RT "Alu alert."
 RL Nature 371:752-752 (1994).
 RN [2]
 RP CONCEPT.
 RX MEDLINE=92241891; PubMed=1572661;
 RA Claverie J.-M.;
 RT "Identifying coding exons by similarity search: alu-derived and other potentially misleading protein sequences."
 RL Genomics 12:838-841 (1992).
 RN [3]
 RP ALU FAMILIES CLASSIFICATION.

RX MEDLINE=88333009; PubMed=3138422;
 RA Quentin Y.;
 RT "The Alu family developed through successive waves of fixation closely connected with primate lineage history."
 RL J. Mol. Evol. 27:194-202 (1988).
 RN [4]
 RP ALU FAMILIES CLASSIFICATION.
 RX MEDLINE=91178815; PubMed=1706781;
 RA Jurka J., Milosavljevic A.;
 RT "Reconstruction and analysis of human Alu genes."
 RL J. Mol. Evol. 32:105-121 (1991).
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning consensus sequences have been constituted that contain all six frames conceptual translations of each of these classes of Alu repeats.
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop codon, 'XXX' is used to separate the various translation phases.
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further pollution of protein sequence databases with Alu-derived amino acid sequences.
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and primate genomes with an average spacing of 4 kb. Some of them are actively transcribed by pol III. Normal transcripts may contain Alu-derived sequences in 5' or 3' untranslated regions. However, cDNA libraries also contain partial and/or rearranged cDNAs ligated with Alu-derived sequence in any orientation. Although Alu elements (especially situated on the complementary strand) have a great potential to create additional/alternative exons, consideration should be given to the possibility that the presence of an Alu in an open reading frame may have resulted from a cloning artifact or may be due to misinterpretation of sequencing data. This point has been overlooked on several occasions, with the consequence of erroneous Alu-derived amino acid sequences being reported.
 CC -!- CAUTION: Any significant similarity of a putative protein sequence with an Alu-translated entry must be taken as a warning that a part of Alu repeat may have been artifactually included in the coding nucleotide sequence.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; U14570; -; NOT_ANNOTATED_CDS.
 KW Hypothetical protein.
 FT DOMAIN 1 98 Frame-1.
 FT DOMAIN 102 199 Frame-2.
 FT DOMAIN 202 300 Frame-3.
 FT DOMAIN 304 401 Frame-4.
 FT DOMAIN 405 502 Frame-5.
 FT DOMAIN 506 603 Frame-6.
 SQ SEQUENCE 603 AA; 65272 MW; B8AAD0AD46BEA114 CRC64;
 Query Match 17.5%; Score 356.5; DB 1; Length 603;
 Best Local Similarity 30.3%; Pred. No. 7.3e-23;
 Matches 112; Conservative 28; Mismatches 115; Indels 115; Gaps 14;
 QY 76 PSVSASQSAARYRTGHHARLCLANFCGRRNRLVSLMCPGWSPELKOSTCLSLPKCWDYRRAAV 135
 D 156 PSYGGXGGRMAXTREAEALVSRDCA---TAVRSPAWATERDSVS-----KXXXX 202
 QY 136 PGLFILFLRHRC-PTLTQDEVQW-----DHSSLQSPSTP----- 169
 D 203 PGA-----VAHACNPSTLGRGGWIMRSGDRDPHGQXGQETPSLLIKIQISRARWRAPVVP 257
 QY 170 -----ELKHP-----PASASQVAGTKDMHHYTWLFIIFIFLFLR 203
 D 258 ATGEAEAGEWRPEPKRSLQXAEIAPLQSAVRPGRQSETPSQKXXF-----FLR 307


```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46542.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K.,
RA Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128399; BAC87422.1; -.
SQ SEQUENCE 176 AA; 19060 MW; 98D1CAC1940AD8FA CRC64;

Query Match 16.3%; Score 332; DB 2; Length 176;
Best Local Similarity 47.9%; Pred. No. 2.7e-21;
Matches 90; Conservative 11; Mismatches 51; Indels 36; Gaps 8;

Qy 208 SVTQAGVQWNLGSLQPLPPGKFLFSCPLSSWDYRRPPRLAN-----FF 253
Dy 5 SVSQAGVQWRD-----PPAHCIL---HLSSSD---SPASASRVAVISGTCHHAQLIF 51

Qy 254 VFLVEMGFTWPA-----RLILISGCDLPASASAGITGVSHARLIFNP-CLFEMESH 308
Dy 52 VFLVETGFCYVQAGRLTLTSSDP---PTSASLSAGIAGLSHRARPFCLFCLVFTVSHS 108

Qy 309 VTQAGVQWNLGSLQPLPPGKFLFSCPLSSWDYRHPPHPAN-FCIFIRGGVSPVLSG 367
Dy 109 VTQGMQWCDHSSLSQSRPPGLKRSPLSSWDYRQASPLANYFFFFIRGKASLCCPG 168

Qy 368 WSQTPDLR 375
Dy 169 WSQTPELK 176

RESULT 13
Q6ZVI4 PRELIMINARY; PRT; 170 AA.
AC Q6ZVI4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ42545.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK124536; BAC85877.1; -.
SQ SEQUENCE 170 AA; 19351 MW; A019980274CEA078 CRC64;

Query Match 16.3%; Score 331; DB 2; Length 170;
Best Local Similarity 42.0%; Pred. No. 3.1e-21;
Matches 71; Conservative 16; Mismatches 39; Indels 44; Gaps 3;

```

```

Qy 251 NFFVFLVEMGFTWPAFL-----ILISGCDLPASASAGITG-----VSH 291
Dy 2 NRLFLLFKSQGLALLPRLKSCGAIHAHCNFFELGSSNFPDILASBRAGTAFGTVVLIIRLSN 61
Qy 292 HARLIENF-----CLFEMESHVTOAGVQWNLGSLQPLP 326
Dy 62 HIAMLDWFWRKQTIWSTRTLNHHHLVSCISFIIIFETESHVTOAGVQWNLSSLOPPP 121

Qy 327 PGLKRFSCLSLSPSSWDYGHLPHPHPANFCIFIRGGVSPVLSGWSQTPDLR 375
Dy 122 PGKRFSCSLTPTSDYMQMPPCLANFCIFSRDGVSPYWPQWRSRTDPCR 170

RESULT 14
ALU7 HUMAN
ID ALU7 HUMAN STANDARD; PRT; 593 AA.
AC P39194;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alu subfamily SQ sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Claverie J.-M., Makalowski W.;
RT "Alu alert.";
RL Nature 371:752-752 (1994).
RN [2]
RP CONCEPT.
RA Claverie J.-M.;
RL MEDLINE=92241891; PubMed=1572661;
RT "Identifying coding exons by similarity search: alu-derived and other
potentially misleading protein sequences.";
RL Genomics 12:838-841 (1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RA Quentin Y.;
RL MEDLINE=88333009; PubMed=3138422;
RT "The Alu family developed through successive waves of fixation closely
connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202 (1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RA MEDLINE=91178815; PubMed=1706781;
RA Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121 (1991).
CC -1- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
consensus sequences have been constituted that contain all six
frames conceptual translations of each of these classes of Alu
repeats.
CC -1- MISCELLANEOUS: Isolated 'x' indicates the presence of a stop
codon, 'xxx' is used to separate the various translation phases.
CC -1- CAUTION: This Alu entry is provided in order to avoid the further
pollution of protein sequence databases with Alu-derived amino
acid sequences.
CC -1- CAUTION: Alu repetitive sequences are interspersed in human and
primate genomes with an average spacing of 4 kb. Some of them are
actively transcribed by pol III. Normal transcripts may contain
Alu-derived sequences in 5' or 3' untranslated regions. However,
cDNA libraries also contain partial and/or rearranged cDNAs
ligated with Alu-derived sequence in any orientation. Although Alu
elements (especially situated on the complementary strand) have a
great potential to create additional/alternative exons,
consideration should be given to the possibility that the presence
of an Alu in an open reading frame may have resulted from a
cloning artifact or may be due to misinterpretation of sequencing

```

CC data. This point has been overlooked on several occasions, with
CC the consequence of erroneous Alu-derived amino acid sequences
CC being reported.
CC -!- CAUTION: Any significant similarity of a putative protein sequence
CC with an Alu-translated entry must be taken as a warning that a
CC part of Alu repeat may have been artifactually included in the
CC coding nucleotide sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14573; -; NOT ANNOTATED CDS.
DR ProdDom; PD003738; GIDA; 1.
KW Hypothetical protein.
FT DOMAIN 1 97 Frame-1.
FT DOMAIN 101 196 Frame-2.
FT DOMAIN 200 295 Frame-3.
FT DOMAIN 299 395 Frame-4.
FT DOMAIN 399 494 Frame-5.
FT DOMAIN 498 593 Frame-6.
SQ SEQUENCE 593 AA; 64417 MW; 54A4F50F33A6089F CRC64;

Query Match 16.2%; Score 330.5; DB 1; Length 593;
Best Local Similarity 45.9%; Pred. No. 1.4e-20;
Matches 89; Conservative 9; Mismatches 53; Indels 43; Gaps 7;

Qy 4 SLLPRLCNGAISAHNRNLRPGSSDSPASAPVAGITGCTHARLILYFFLVEMEFLLHV 63
Db 405 SLLPRLCNGAISAHNRNLRPGSSDSPASAPVAGITGCTHARLILYFFLVEMEFLLHV 463

Qy 64 GQAGLELPTSDPSVSASQSAARYRTG--HHARLCLANFCGRNRYSLMCPSPWS----- 113
Db 464 GQAGLELPTSDPSVSASQSAARYRTG--HHARLCLANFCGRNRYSLMCPSPWS----- 113

Qy 114 -----PELKQSTCLSLPKCWYRAAIVPGLFILFLRHRCPTLTQDEVQWCDH 161
Db 520 ATASRVOAILLPQPEXGLQAP-----ATFGXFLYFXRRGFTMLAR----- 564

Qy 162 SSLOPSTPEIKHPP 175
Db 565 -LVNSXFPQVIHPP 577

RESULT 15
Q6ZSJ7 PRELIMINARY; PRT; 143 AA.
AC Q6ZSJ7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ45460.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Subthalamic nucleus;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK127382; BAC6951.1; -;
SQ SEQUENCE 143 AA; 15771 MW; AD4FBFB10F6075C CRC64;

Query Match 16.0%; Score 325; DB 2; Length 143;
Best Local Similarity 37.5%; Pred. No. 8.7e-21;
Matches 90; Conservative 9; Mismatches 39; Indels 102; Gaps 5;

Qy 3 FSLLPRLCNGAISAHNRNLRPGSSDSPASAPVAGITGCTHARLILYFFLVEMEFLLHV 62
Db 6 FLSLSPKLECSGAILAHFNLCLLGSSDSPASAPVAGITGCTHARLILYFFLVEMEFLLHV 63

Qy 63 VGOAGLELPTSDPSVSASQSAARYRTGHHARLCLANFCGRNRYSLMCPSPWSPELKQSTCL 122
Db 64 VGOAGLELPTSDPSVSASQSAARYRTGHHARLCLANFCGRNRYSLMCPSPWSPELKQSTCL 122

Qy 123 SLPKCWYRAAIVPGLFILFLRHRCPTLTQDEVQWCDHSLQPSLPSTPEIKHPPASASQVA 182
Db 77 -----LPASASQSA 85

Qy 183 GTKDMHHYTWLIFIFIPNLRQSLNSVTQAGVQWRNIGSLQPLPPGPKLFSCPSLLSSWD 242
Db 86 GITGVSHCTWAFPPFFFLRQSF-IVIEAGVQWRDLSLQPLPPGPKLFSCPSLLSSWD 143

Search completed: September 8, 2005, 10:44:16
Job time : 378 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 09:57:54 ; Search time 101 Seconds
(without alignments)
277.162 Million cell updates/sec

Title: US-10-092-934-10

Perfect score: 2034

Sequence: 1 MEFSLLPLRLECGAISAHR.....FIRGGVSPVLSGWSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	375	2	US-08-454-557C-121
2	2034	100.0	375	2	US-08-340-426D-121
3	2034	100.0	375	2	US-08-450-673C-121
4	2034	100.0	375	5	US-09-872-968-2
5	1415.5	69.6	397	5	PCT-US95-17111A-121
6	267.5	13.2	96	4	US-09-513-999C-6065
7	259.5	12.8	144	4	US-09-513-999C-6953
8	256	12.6	108	4	US-09-513-999C-7878
9	254.5	12.5	239	4	US-09-800-729-193
10	254.5	12.5	310	4	US-09-800-729-190
11	253	12.4	118	4	US-09-663-600A-114
12	251	12.3	776	4	US-10-020-079-24
13	251	12.3	789	4	US-10-020-079-22
14	251	12.3	863	4	US-10-020-079-32
15	251	12.3	876	4	US-10-020-079-30
16	251	12.3	889	4	US-10-020-079-20
17	251	12.3	895	4	US-10-020-079-18
18	251	12.3	976	4	US-10-020-079-28
19	251	12.3	982	4	US-10-020-079-26
20	250	12.3	97	4	US-09-513-999C-4770
21	250	12.3	132	4	US-09-636-215-573
22	250	12.3	132	4	US-09-685-166A-573
23	250	12.3	132	4	US-09-679-426-573
24	250	12.3	132	4	US-09-759-143-573
25	250	12.3	132	4	US-09-651-236-573
26	250	12.3	135	4	US-09-685-166A-884
27	250	12.3	135	4	US-09-679-426-884

28 250 12.3 135 4 US-09-759-143-884 Sequence 884, App
29 247.5 12.2 112 4 US-09-513-999C-7870 Sequence 7870, App
30 246 12.1 119 4 US-09-513-999C-7867 Sequence 7867, App
31 245 12.0 102 4 US-09-621-976-6112 Sequence 6112, App
32 242.5 11.9 121 4 US-09-513-999C-7874 Sequence 7874, App
33 241 11.8 91 4 US-09-621-976-5929 Sequence 5929, App
34 239 11.3 84 4 US-09-621-976-7167 Sequence 7167, App
35 236 11.1 61 4 US-09-513-999C-4581 Sequence 4581, App
36 224.5 11.0 396 4 US-09-949-016-6783 Sequence 6783, App
37 224.5 11.0 442 4 US-09-949-016-6762 Sequence 6762, App
38 223.5 11.0 87 4 US-09-621-976-5968 Sequence 5968, App
39 219 10.8 76 4 US-09-621-976-6338 Sequence 6338, App
40 218 10.7 1079 3 US-09-058-489-22 Sequence 22, Appl
41 217.5 10.7 87 4 US-09-205-258-342 Sequence 342, App
42 216 10.6 169 4 US-09-663-600A-208 Sequence 208, App
43 215.5 10.6 108 4 US-09-513-999C-4433 Sequence 4433, App
44 214.5 10.5 83 4 US-09-621-976-5396 Sequence 5396, App
45 211.5 10.4 103 4 US-09-513-999C-5327 Sequence 5327, App

ALIGNMENTS

RESULT 1

US-08-454-557C-121
; Sequence 121, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-454-557C-121

Query Match 100.0%; Score 2034; DB 2; Length 375;

Best Local Similarity 100.0%; Pred. No. 6.2e-215;

Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECGAISAHRNLRLPGSSSDSPASPVAGITGCTHARLLYFLVEMEF 60

Db 1 MEFSLLPLRLECGAISAHRNLRLPGSSSDSPASPVAGITGCTHARLLYFLVEMEF 60

Qy 61 LHVQAGLELPTSDPVSASQSAARYNTGHARLCLANFCGRNRVSLMCPSPWSPDLQST 120

```
Db 1 LHVGAGLEPTSDPVSASQARVYRGHARLCLANFCGRNVSVMCPSPWSPKQST 120
Qy 121 CLSLPKCDWYRAAVPGLFILFLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Db 121 CLSLPKCDWYRAAVPGLFILFLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTOAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTOAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

RESULT 2

```
US-08-340-426D-121
; Sequence 121, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wanda, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-340-426D-121
```

```
Query Match 100.0%; Score 2034; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 6.2e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
```

```
Qy 61 LHVGAGLEPTSDPVSASQARVYRGHARLCLANFCGRNVSVMCPSPWSPKQST 120
Db 61 LHVGAGLEPTSDPVSASQARVYRGHARLCLANFCGRNVSVMCPSPWSPKQST 120
Qy 121 CLSLPKCDWYRAAVPGLFILFLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Db 121 CLSLPKCDWYRAAVPGLFILFLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTOAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTOAGVQWNRNLGSLQPLPGFKLFSCPSLLSS 240
Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGSLOPLPGKRFKFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

RESULT 3

```
US-08-450-673C-121
; Sequence 121, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wanda, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450.673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-450-673C-121
```

```
Query Match 100.0%; Score 2034; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 6.2e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60
```

QY 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
DB 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
QY 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
DB 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTOAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTOAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
QY 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVTOAGVQWRNLGSLQPLPPGLKRFCSLSLPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVTOAGVQWRNLGSLQPLPPGLKRFCSLSLPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375
RESULT 4
US-09-872-968-2
; Sequence 2, Application US/09872968
; Patent No. 6770797
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/09/872, 968
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-872-968-2

Query Match 100.0%; Score 2034; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 6.2e-215;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
DB 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
QY 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
DB 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCPSPKQST 120
QY 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
DB 121 CLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTOAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTOAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
QY 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVTOAGVQWRNLGSLQPLPPGLKRFCSLSLPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVTOAGVQWRNLGSLQPLPPGLKRFCSLSLPSSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375
RESULT 5
PCT-US95-17111A-121
; Sequence 121, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-17111A-121

Query Match 69.6%; Score 1415.5; DB 5; Length 397;
Best Local Similarity 74.6%; Pred. No. 6.8e-147;
Matches 285; Conservative 10; Mismatches 44; Indels 43; Gaps 7;
QY 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
DB 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
QY 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCP 110
DB 61 LHVQAGLELPTSDPSVASQARVYTGHHARLCLANFCGRNRVSLMCP 110
QY 111 SMSPELKQSTCLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPE 170
DB 111 SMSPELKQSTCLSLPKCWDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPE 170
QY 171 ----IKHPASASQVAGTKDMHYTWLIFIFNFRQSLNSVTOAGVQWRNLGSLQPLP 226
DB 171 SLLPQPP---KVAGTKDMHYTWLIFIFNFRQSLNSVTOAGVQWRNLGSLQPLP 226
QY 227 PGKLFSCPSLLSSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 278
DB 227 PGKLFSCPSLLSSWDYRRPRLANFFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 278
QY 279 SASQSAGITGVSHHARLIFNFCIFEMESHVTOAGVQWRNLGSLQPLPPGKLFSCPSLLSS 338

```

Db      282  --PKVLGLQDVTPTARPIFNCLFEMESHVSQTQGVQWPNLGLSLQPLPPGLKRFSCLSLP 339
Qy      339  SSWDYGHLPHPANFCIRGG 360
Db      340  SSWDYGHLPHTPLIFVFSLEAG 361

```

```

RESULT 6
US-09-513-999C-6065
; Sequence 6065, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6065
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6065

```

```

Query Match      13.2%; Score 267.5; DB 4; Length 96;
Best Local Similarity 60.4%; Pred. No. 1.5e-21;
Matches 61; Conservative 3; Mismatches 22; Indels 15; Gaps 2;

Qy      205  SLNSVTQGVQWNRNGSLQPLPPGFKLPSCPSLLSSWDYRPPRIANFFVFLVMGFTMF 264
Db      2    SFTLVAQGVQWRDGLGSPQLPPGFKRFSCFLSPSSWDYRHVPLLPANLFLVEMGF--- 58
Qy      265  ARLLILISGPC-----DLPASASQSAGITGVSHHARLIF 297
Db      59  ----LHVGPAGLGLPSTGDLFPASASQSAGIIGVSHRSQPF 95

```

```

Query Match      13.2%; Score 267.5; DB 4; Length 96;
Best Local Similarity 60.4%; Pred. No. 1.5e-21;
Matches 61; Conservative 3; Mismatches 22; Indels 15; Gaps 2;

Qy      205  SLNSVTQGVQWNRNGSLQPLPPGFKLPSCPSLLSSWDYRPPRIANFFVFLVMGFTMF 264
Db      2    SFTLVAQGVQWRDGLGSPQLPPGFKRFSCFLSPSSWDYRHVPLLPANLFLVEMGF--- 58
Qy      265  ARLLILISGPC-----DLPASASQSAGITGVSHHARLIF 297
Db      59  ----LHVGPAGLGLPSTGDLFPASASQSAGIIGVSHRSQPF 95

```

```

RESULT 7
US-09-513-999C-6953
; Sequence 6953, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6953
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 9
; OTHER INFORMATION: Xaa= * or Cys or Trp
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 10
; OTHER INFORMATION: Xaa=Leu or Met or Val

```

```

; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 72
; OTHER INFORMATION: Xaa=Glu or Val
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 110
; OTHER INFORMATION: Xaa=Asp or His or Asn or Tyr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 132
; OTHER INFORMATION: Xaa=Asn or Thr
US-09-513-999C-6953

```

```

Query Match      12.8%; Score 259.5; DB 4; Length 144;
Best Local Similarity 72.0%; Pred. No. 2e-20;
Matches 59; Conservative 3; Mismatches 19; Indels 1; Gaps 1;

Qy      3    FSLLLPRLECNAGISAHRNLRLEFGSSDSPASASPVAGITGMCTHARLILYFFLVEMEFLLH 62
Db      23  FSLLSRPECNGVILAHCNLRLLSGSTHSPVSASGVAGITGMHHAQLI-FXFLVETGFFH 81
Qy      63  VQAGLELPTSDDDPSVSASQSA 84
Db      82  VQAGLELLTSGDPPASASQSA 103

```

```

RESULT 8
US-09-513-999C-7878
; Sequence 7878, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7878
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -17..-1
; OTHER INFORMATION: score 5.7
; OTHER INFORMATION: seq VVCCSCSWFFLFC/FV
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 56
; OTHER INFORMATION: Xaa= * or Cys or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 59
; OTHER INFORMATION: Xaa=Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 61
; OTHER INFORMATION: Xaa=Asp or Asn
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 62
; OTHER INFORMATION: Xaa=Ala or Glu or Gly or Lys or Met or Arg or Thr or Val
; NAME/KEY: UNSURE
; LOCATION: 64
; OTHER INFORMATION: Xaa= * or Leu

```



```

; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 66
; OTHER INFORMATION: Xaa=Asn or Tyr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 67
; OTHER INFORMATION: Xaa= * or Tyr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 71
; OTHER INFORMATION: Xaa=Pro or Ser
US-09-513-999C-7878

Query Match      12.6%; Score 256; DB 4; Length 108;
Best Local Similarity 66.7%; Pred. No. 3.2e-20;
Matches 54; Conservative 3; Mismatches 20; Indels 4; Gaps 3;

Qy 297 FNFC-LFEMESHVTOAGVQWPNLGSLOPLPGLKRFSCLSLPSSWDYGHLPHPAN--F 353
Db 14 FLFCFVEMESHVTOAGVQWPNLGSLOPLPGLKRFSCLSLPSSWDYRCLPPSSANFFF 72

Qy 354 CIFIRGGVSPYLSGWSQTPDL 374
Db 73 XIFXXXVXXPGWQXSPDL 93

RESULT 9
US-09-800-729-193
; Sequence 193, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 193
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-193

Query Match      12.5%; Score 254.5; DB 4; Length 239;
Best Local Similarity 64.2%; Pred. No. 1.5e-19;
Matches 61; Conservative 6; Mismatches 25; Indels 3; Gaps 3;

Qy 1 MEFSLLL-PRLECNAGISAHRNLRPGSSDSPASPVAGITGMCTHARLILYFFLVEME 59
Db 120 LRWSLTLSPRLECCSAISAHCNLRPGSSNSPALASQVAGITGICHARQI-FVFLVETG 178

Qy 60 FLHVQAGLELPTSDPDSVSASQAR-YRTGHAR 93
Db 179 FCHVQAGLELLISGDSPPASAFQSGAGITGVSHR 213

RESULT 10
US-09-800-729-190
; Sequence 190, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013

```

```

; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 190
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-190

Query Match      12.5%; Score 254.5; DB 4; Length 310;
Best Local Similarity 64.2%; Pred. No. 2.2e-19;
Matches 61; Conservative 6; Mismatches 25; Indels 3; Gaps 3;

Qy 1 MEFSLLL-PRLECNAGISAHRNLRPGSSDSPASPVAGITGMCTHARLILYFFLVEME 59
Db 191 LRWSLTLSPRLECCSAISAHCNLRPGSSNSPALASQVAGITGICHARQI-FVFLVETG 249

Qy 60 FLHVQAGLELPTSDPDSVSASQAR-YRTGHAR 93
Db 250 FCHVQAGLELLISGDSPPASAFQSGAGITGVSHR 284

RESULT 11
US-09-663-600A-114
; Sequence 114, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 114
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -83...-1
; NAME/KEY: UNSURE
; LOCATION: 28,32
; OTHER INFORMATION: Xaa = any one of the twenty amino acids
US-09-663-600A-114

Query Match      12.4%; Score 253; DB 4; Length 118;
Best Local Similarity 65.9%; Pred. No. 7.9e-20;
Matches 60; Conservative 3; Mismatches 24; Indels 4; Gaps 3;

Qy 204 QSLNSVTOAGVQWPNLGSLOPLPGLKRFSCLSLPSSWDYRPPRL-ANPFVFLVEMGFT 262
Db 5 QSFTLVAQAGVQWPNLGSLOPLPGLKRFSCLSLPSSWDYRPPRPAGFVFLVETGLH 64

Qy 263 MF--ARLILISGPCDLPASASQSGAGITGVSH 291

```


APPLICANT: Friddle, Carl Johan
TITLE OF INVENTION: No. 6579710e1 Human Kinases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0281-USA
CURRENT APPLICATION NUMBER: US/10/020,079
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255,103
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: US 60/289,422
PRIOR FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 876
TYPE: PRT
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(876)
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-020-079-30

Query Match 12.3%; Score 251; DB 4; Length 876;
Best Local Similarity 47.9%; Pred. No. 2.4e-18;
Matches 69; Conservative 9; Mismatches 38; Indels 28; Gaps 8;
Qy 211 QAGVQWRNLGSLQPLPPGFKLFSCPSLLSSWDYRR--PPRLANFFVFLVEMGFT-MPEARL 267
Db 749 QAGVQWRDLGSLQPPPPRFKQFSCSLPSRWDYRHAPPPHPPAN-FVFLVETGFLHVEAGL 807
Qy 268 -ILISGCDLPASASQAGITGVSHHARLIENFCLFEMESHVSVTQAGVQWPNLGSLOPLP 326
Db 808 ELTSG--DLPASASQIAGITGVSHRAQP--EVCEFN-RKHTGQR----- 847
Qy 327 PGLKRFSCSLSPSSWDYGHLPHP 350
Db 848 ---EQMVCAGSERAWMRDLPGRP 868

Search completed: September 8, 2005, 10:47:21
Job time : 102 secs

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 8, 2005, 10:44:23 ; Search time 389 Seconds
(without alignments)
379.650 Million cell updates/sec

Title: US-10-092-934-10

Perfect score: 2034

Sequence: 1 MEPSLLPLRLCNCAISAHRR.....FIRGGVSPYLSCGWSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	375	9	US-09-964-666-2
2	2034	100.0	375	9	US-09-964-412-2
3	2034	100.0	375	10	US-09-964-667-2
4	2034	100.0	375	10	US-09-872-968-2
5	2034	100.0	375	10	US-09-964-678A-2
6	2034	100.0	375	14	US-10-146-130-2
7	2034	100.0	375	14	US-10-092-934-10
8	2034	100.0	375	14	US-10-153-334-1
9	2034	100.0	375	14	US-10-198-069-1
10	2034	100.0	375	14	US-10-157-031-299
11	2034	100.0	375	14	US-10-198-070-1

12	2034	100.0	375	16	US-10-755-889-410	Sequence 410, App
13	2034	100.0	375	17	US-10-910-173-2	Sequence 2, Appli
14	734.5	36.1	361	9	US-09-995-494-107	Sequence 107, App
15	729	35.8	449	14	US-10-007-280A-140	Sequence 140, App
16	569	28.0	241	15	US-10-276-774-1834	Sequence 1834, Ap
17	517	25.4	213	15	US-10-296-115-911	Sequence 911, App
18	471	23.2	286	15	US-10-291-172-654	Sequence 654, App
19	471	23.2	286	15	US-10-221-278-654	Sequence 654, App
20	428.5	21.0	183	9	US-09-989-920-245	Sequence 245, App
21	407	20.0	361	15	US-10-276-774-1862	Sequence 1862, Ap
22	382.5	18.8	341	18	US-10-220-335-586	Sequence 586, App
23	382.5	18.8	673	14	US-10-157-031-291	Sequence 291, App
24	377.5	18.6	158	15	US-10-104-047-3047	Sequence 3047, Ap
25	356.5	17.5	603	16	US-10-408-765A-140	Sequence 140, App
26	353.5	17.4	217	14	US-10-017-161-1956	Sequence 1956, Ap
27	353.5	17.4	217	15	US-10-292-798-1604	Sequence 1604, Ap
28	349	17.2	157	15	US-10-108-260A-4272	Sequence 4272, Ap
29	336	16.5	148	15	US-10-276-774-1784	Sequence 1784, Ap
30	332	16.3	108	14	US-10-078-090-143	Sequence 143, App
31	324	15.9	189	15	US-10-104-047-3196	Sequence 3196, Ap
32	316	15.5	152	14	US-10-198-070-112	Sequence 112, App
33	311	15.3	74	15	US-10-276-774-1929	Sequence 1929, Ap
34	310	15.2	101	15	US-10-094-749-2181	Sequence 2181, Ap
35	309.5	15.2	107	15	US-10-276-774-1874	Sequence 1874, Ap
36	308	15.1	93	10	US-09-764-891-5337	Sequence 5337, Ap
37	308	15.1	93	14	US-10-205-428-486	Sequence 486, App
38	308	15.1	100	13	US-10-016-157A-187	Sequence 187, App
39	307	15.1	114	15	US-10-104-047-2423	Sequence 2423, Ap
40	305.5	15.0	137	15	US-10-276-774-1959	Sequence 1959, Ap
41	303	14.9	203	15	US-10-108-260A-2748	Sequence 2748, Ap
42	302.5	14.9	128	15	US-10-276-774-2194	Sequence 2194, Ap
43	301	14.8	156	15	US-10-104-047-2334	Sequence 2334, Ap
44	300.5	14.8	101	15	US-10-276-774-2026	Sequence 2026, Ap
45	300	14.7	76	15	US-10-276-774-2032	Sequence 2032, Ap

ALIGNMENTS

RESULT 1
US-09-964-666-2
; Sequence 2, Application US/09964666
; Patent No. US20020104108A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,666
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0609.4370000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540


```
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0609.4370000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-667-2

Query Match 100.0%; Score 2034; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVQAGLELPTSDSPSVASQARYTGHARLCLANFCGRNRVSLMCPSPKQST 120
Db 61 LHVQAGLELPTSDSPSVASQARYTGHARLCLANFCGRNRVSLMCPSPKQST 120
Qy 121 CLSLPKCWDYRRAAVPGLFILFLRHCPTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFLRHCPTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Qy 181 VAGTKDMHYTWLIFIFNLRQSLNSVTOAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNLRQSLNSVTOAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
Qy 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
Qy 301 LFEMESHVSVOAGVQWRNLGSLQPLPPGFKLFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVSVOAGVQWRNLGSLQPLPPGFKLFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 5
US-09-964-678A-2
; Sequence 2, Application US/09964678A
; Publication No. US20030066097A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for Screening Drugs
; TITLE OF INVENTION: Effective for the Treatment or Prevention of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4370002
; CURRENT APPLICATION NUMBER: US/09/964,678A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 09/380,203
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US98/03685
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/038,908
; PRIOR FILING DATE: 1997-02-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AD7c-NTP cDNA
US-09-964-678A-2

Query Match 100.0%; Score 2034; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVQAGLELPTSDSPSVASQARYTGHARLCLANFCGRNRVSLMCPSPKQST 120
Db 61 LHVQAGLELPTSDSPSVASQARYTGHARLCLANFCGRNRVSLMCPSPKQST 120
Qy 121 CLSLPKCWDYRRAAVPGLFILFLRHCPTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFLRHCPTLTQDEVQWCDHSSLPQSTPEIKHPPASQ 180
Qy 181 VAGTKDMHYTWLIFIFNLRQSLNSVTOAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNLRQSLNSVTOAGVQWRNLGSLQPLPPGFKLFSCPSLLSS 240

RESULT 4
US-09-872-968-2
; Sequence 2, Application US/09872968
; Publication No. US20030050262A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/09/872,968
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-872-968-2

Query Match 100.0%; Score 2034; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNAGISAHRNLRPLGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
```

Db 181 VAGTKDMHHYTLWIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Qy 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
Db 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 6

US-10-146-130-2
; Sequence 2, Application US/10146130
; Publication No. US20030004107A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
; FILE REFERENCE: NEURAL THREAD PROTEINS
; FILE REFERENCE: 59003.000007
; CURRENT APPLICATION NUMBER: US/10/146,130
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-130-2

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVGQAGLELPTSDPVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Qy 121 CLSLPKCDYRRAAVPGLFILFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
Db 121 CLSLPKCDYRRAAVPGLFILFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
Qy 181 VAGTKDMHHYTLWIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Db 181 VAGTKDMHHYTLWIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Qy 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
Db 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 7

US-10-092-934-10
; Sequence 10, Application US/10092934
; Publication No. US20030054990A1
; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 018792-0199
; CURRENT APPLICATION NUMBER: US/10/092,934
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/273,957
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-934-10

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVGQAGLELPTSDPVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Qy 121 CLSLPKCDYRRAAVPGLFILFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
Db 121 CLSLPKCDYRRAAVPGLFILFLRHRCTTLTQDEVQWCDHSSLPQSTPEIKHPPASASQ 180
Qy 181 VAGTKDMHHYTLWIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Db 181 VAGTKDMHHYTLWIFIFNFRLRQSLNSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 240
Qy 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSAGITGVSHHARLIENFC 300
Db 241 WDYRRPPLANFFVFLVEMGTFMFARLLIISGPCDLPASASQSAGITGVSHHARLIENFC 300
Qy 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
Db 301 LFEMESHVSVTQAGVQWRNLGSLQPLPGFKLFCPSLLSS 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 8

US-10-153-334-1
; Sequence 1, Application US/10153334
; Publication No. US20030096350A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003-000006
; CURRENT APPLICATION NUMBER: US/10/153,334
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,156
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-334-1
Query Match 100.0%; Score 2034; DB 14; Length 375;

Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNCAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFLVEMEF 60
Db 1 MEFSLLPLRLECNCAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKOST 120
Db 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 9
US-10-198-069-1
; Sequence 1, Application US/10198069
; Publication No. US20030096756A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000009
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: US/10/198,069
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-198-069-1

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNCAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFLVEMEF 60
Db 1 MEFSLLPLRLECNCAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKOST 120
Db 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 10
US-10-157-031-299
; Sequence 299, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 299
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-157-031-299

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLPLRLECNCAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFLVEMEF 60
Db 1 MEFSLLPLRLECNCAISAHRNLRPGSSDSPASASPVAGITGCTHARLILYFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKOST 120
Db 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPWSPKOST 120

Qy 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDWYRRAAVPGLFILFRLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAAGITGVSHHARLIENFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQSAAGITGVSHHARLIENFC 300

Qy 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTOAGVQWPNLGLQPLPGPKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

```
RESULT 11
US-10-198-070-1
; Sequence 1, Application US/10198070
; Publication No. US20030109437A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-1

Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120

Qy 121 CLSLPKCWDYRRAAVPGLFILFHLRHCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFHLRHCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300

Qy 301 LFEMESHVTVQAGVQWRNLSLQPLPGFKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTVQAGVQWRNLSLQPLPGFKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 12
US-10-755-889-410
; Sequence 410, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
```

```
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 410
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-410

Query Match 100.0%; Score 2034; DB 16; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120

Qy 121 CLSLPKCWDYRRAAVPGLFILFHLRHCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFILFHLRHCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180

Qy 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLLSS 240
Db 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLSLQPLPGFKLFSCPSLLSS 240

Qy 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQASAGITGVSHHARLIFNFC 300

Qy 301 LFEMESHVTVQAGVQWRNLSLQPLPGFKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
Db 301 LFEMESHVTVQAGVQWRNLSLQPLPGFKLRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360

Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 13
US-10-910-173-2
; Sequence 2, Application US/10910173
; Publication No. US2005009041A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/10/910,173
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/872,968
; PRIOR FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-910-173-2

Query Match 100.0%; Score 2034; DB 17; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

Qy 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQSAARYRTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120
```

```
Db 61 LHVGAGLELPTDDPSVSASQARYTGHARLCLANFCGRNRVSLMCPFSWSPKLQST 120
Qy 121 CLSLPKCWDYRRAAVPGLFLFLRHRCPRTLQDEVQWCHDSSLQSPSTPEIKHPPASASQ 180
Db 121 CLSLPKCWDYRRAAVPGLFLFLRHRCPRTLQDEVQWCHDSSLQSPSTPEIKHPPASASQ 180
Qy 181 VAGTKDMHHTWLIPIFIFNLRQSLNSVTQAGVQWNRNLGSLQPLPPGKLFSCPSLLSS 240
Db 181 VAGTKDMHHTWLIPIFIFNLRQSLNSVTQAGVQWNRNLGSLQPLPPGKLFSCPSLLSS 240
Qy 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGBCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPRLANFFVFLVEMGFTMPARLILISGBCDLPASASQAGITGVSHHARLIFNFC 300
Qy 301 LFEMESHVSVTQAGVQWNRNLGSLQPLPPGKLFSCPSLLSSWDYCHLPPHPANFCIFIRGG 360
Db 301 LFEMESHVSVTQAGVQWNRNLGSLQPLPPGKLFSCPSLLSSWDYCHLPPHPANFCIFIRGG 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375

RESULT 14
US-09-995-494-107
; Sequence 107, Application US/09995494
; Patent No. US20020127578A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Chen, Sei-Yu
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0293
; CURRENT APPLICATION NUMBER: US/09/995,494
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/253,176
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 107
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-995-494-107

Query Match 36.1%; Score 734.5; DB 9; Length 361;
Best Local Similarity 43.1%; Pred. No. 4.3e-61;
Matches 185; Conservative 24; Mismatches 93; Indels 127; Gaps 12;

Qy 6 LLPLRCNGAISAHNRNLRLPGSSDPSASQSPVAGITGM-----CTHARL----- 49
Db 1 LLPLECSGITHMAHCKLLKLGSLDLPASARVGGTGMRTGRTMGSTCAHFQIGLFPVAN 60
Qy 50 -----ILYFLVEMEFHVGQAGLELPTDDPSVSASQARYTGHARLCL 97
Db 61 FERSFYILGTSPLDLMLINMYF-----FALLIHVLF- 94
Qy 98 NPGCRNRVSLMCP--SNWSPKQSTCLSLPKCWDYRRAAVPGLFILPFLRHC----- 148
Db 95 -----NRDLSLCCPGASUTGLGEAFACLSLPKCWDYTYGPLL-LPIDIFLQVCCLSFYFLN 148
Qy 149 PTLTQDEVQWCHDSSLQSPSTPEIKHPPASASQAGTKDMHHTWLIPIFIFNLRQSLNS 208
Db 149 TVFQRAEV-----LIFFQFVYLRLWSLHS 172
Qy 209 VTQAGVQWRNLGSLQPLPPGKLFSCPSLLSSWDYRR--PPRLAN--FFVFLVEMGFTMFA 265
```

```
Db 173 VAQAGVQWNLGSLQPLPPRFRFRFSCLSLLSSWDHRHAPCLANFLFFFKELVDQSFMTLA 232
Qy 266 RLILISGPC-DLPASASQAGITGVSH-----HARLIFNCLFEMES 306
Db 233 RLVLNSAPSGDLPPAPASQAGITGVHRCTWPKSFKPADSHIGLAFHFAFFFFFAVAS 292
Qy 307 HSVTQAGVQWNLGSLQPLPPGKLFSCLSLSSWDYCHLPPHPANFCIFIRGVSPLYLS 366
Db 293 HPTAQAGVQWNRDLGSLQPPPGKFLCLSLPGSWDYRRAPPRQANFCIPSRDGVSPCWT 352
Qy 367 GWSQTPDLR 375
Db 353 GWSQTPDLR 361

RESULT 15
US-10-007-280A-140
; Sequence 140, Application US/10007280A
; Publication No. US20030059784A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Chenghua, Liu
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; FILE REFERENCE: DEX-0257
; CURRENT APPLICATION NUMBER: US/10/007,280A
; CURRENT FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,640
; PRIOR FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 140
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-007-280A-140

Query Match 35.8%; Score 729; DB 14; Length 449;
Best Local Similarity 42.6%; Pred. No. 1.9e-60;
Matches 196; Conservative 40; Mismatches 122; Indels 102; Gaps 19;

Qy 6 LLPLRCNGAISAHNRNLRLPGSSDPSASQSPVAGITGMCTHARL-----ILYFLVE 57
Db 1 LFPRLVYGGTILAYCNLHLPGSSNPPTSASQVAGTRDVCHHTWLVCVCVCVCVCVE 60
Qy 58 MEFLHVGQAGLELPTDDPSVSASQARYTGHARL-----LANFC-- 100
Db 61 MRPHYVSQAGLELSSDDPISASQAGIIGISH---CTWPWHDSTFISPGAEPLTFAYTW 117
Qy 101 -GENRVSL-----MCPWSPELKQSTCLSLPKCWDYRR-----AAVFL-----F 139
Db 118 PGRPKPIPTLILLYPGGDVLAFTRELYYASRQPGASDTARESGNAGVDFLHKW 177
Qy 140 ILFFLRHRCPRTLQDEV-----QWCD--HSSLQSPSTPEIKHPPA-----SASQVAGTKD 186
Db 178 LIF-----CPFSNQSHLWTTKSWAEVPHPGRAELPAMKEQAANENSGSVTEPSSAS 232
Qy 187 MHHTWLIPIFIFN-----FLRQSLNSVTQAGVQW-----RNLGSLQPLPPGKLFSCPSLLS 239
Db 233 ILHARWDVY-FLINALIYFLRQSLRVAQAGVQWCGADIGSLQPLPPGKAPFCLLS 291
Qy 240 SWDYRR--PPRLANFFVFLVEMGFTMPARLILISGBCDLPASASQAGITGVSHHAR---- 294
Db 292 SWDYRSLPPCPANFFVFLIETGFHHISQ-ISISAPCDPPASASQAGITGWSHCAQPDVY 350
Qy 295 -----LIFNCLFEMESHVSVTQAGVQWPNLG-----SLQPLPGLKRFSC 334
Db 351 YVSVGYIKQDRCYLFRRFFFTESRTVAQAG-RLERSGAISTRSLQPLPGLKRFSC 409
Qy 335 LSLPSSWDYCHLPPHPANFCIFIRGVSPLYSGWSQTPDL 374
Db 410 LSLSSWDYRCTPFLAHFCTFRSDGVSPCWGWSLSDDL 449
```

Search completed: September 8, 2005, 11:17:14
Job time : 390 secs
